Investigating the spatial distribution of campylobacteriosis in New Zealand

---

A thesis submitted in fulfilment of the requirements for the Degree
of Master of Science in Geography
in the University of Canterbury
by Esther Christiane Rind

University of Canterbury
2007

Abstract

Background

Infection with *Campylobacter* is thought to account for about 5% - 14% of all food and waterborne diarrhoea cases worldwide. By international standards, New Zealand has extremely high rates of campylobacteriosis which are thought to be the highest reported rates worldwide. The incidence has been steadily increasing since 1980 (when the disease became notifiable), reaching a peak of cases in 2003 (396/100,000). Although different surveillance systems complicate international comparisons, New Zealand's particularly high rate still lacks a conclusive explanation.

Aims

This study investigates the geographical distribution of campylobacteriosis in New Zealand and the relative importance of factors assumed to be affecting the distribution of this disease, including those related to climate, landuse, water and food. The approach aims to explain why certain areas might increase the probability of becoming infected.

Methodology

A Geographical Information System (GIS) is used to visualise the disease rate, investigate potential disease clustering and identify outliers. Hierarchical regression, including the analysis of residuals, is applied to analyse the variables in their complex interrelation and to investigate whether there is statistical evidence explaining the geographical variation in campylobacteriosis. This study is undertaken at the territorial local authority level, as all required data are available at this spatial scale and covers the period 1997 to 2005.

Results and conclusion

There is a large geographical variation in campylobacteriosis across New Zealand, ranging from an average annual rate of 97/100,000 to 526/100,000 per territorial local authority (TLA). Generally, there is statistical evidence for global and local clustering of the disease rate. There are upper and lower outliers of campylobacteriosis in New Zealand; however, higher rates primarily appear in the South Island. The hierarchical modelling confirms statistical significance for some of the environmental and sociodemographic variables. The final model explains about 58% of the variation in campylobacteriosis, and the residuals reflect this variation relatively accurately in approximately 75% of all TLAs. Although the evaluation of the results is confronted with a number of challenges, it is concluded that socioeconomic and demographic factors are crucial factors in explaining the observed spatial patterns in the notification data.

Acknowledgements

Writing this thesis in New Zealand happened to be quite by chance. My former supervisor PD Dr. Thomas Kistemann from the University of Bonn in Germany asked me if I wanted to apply for a scholarship in Christchurch which I firstly assumed to be in England. It turned out the travelling was a little bit more time consuming, from a European point of view literally to the end of the world. However, many wonderful opportunities resulted from Thomas' initial idea, not only of academic nature, but also including new friendships and some travelling in an amazing country.

The scholarship granted from the GeoHealth Laboratory provided the financial basis for this year and also included an informative and excellent internship at the Ministry of Health in Wellington. Thanks to Dr. Jamie Pearce and Dr. Paul White, directors of the GeoHealth Laboratory, for their input and the time they took discussing future ideas and plans. Moreover, they afforded the opportunity to give my first talk in English at the GeoHealth Conference in Nelson, remembered as a valuable and encouraging experience.

My supervisors, Dr. Jamie Pearce and Dr. Clive Sabel, offered ongoing help and direction throughout the year. Most notably, the continuous advice concerned with the precise and clear formulation of ideas and the introduction to the implementation of multiple regression was very much appreciated.

Thanks to the warm welcome from everybody working in GeoHealth Laboratory, including some memorable leisure activities, which made my year in Christchurch very easy and lively. Moreover, a myriad of questions, particularly starting with "How do you say that in English?" were always answered with friendliness and patience. A very special thank you to Laura Miller for statistical, formal and linguistic advice, not to mention all of our coffee, tea and food breaks.

In terms of references, personal discussions or contact via email, Dr. Michael Baker, Dr. Iain Lake, Erin Holmes, Sarah Cook and Neil Andrew de Wet offered valuable advice as their previous work is directly related to this project.

Carole Acheson undertook the time consuming task of reading the complete thesis for the first time. Her linguistic improvements were very much appreciated, and she offered her support free of charge.

The regular contact with my siblings as well as my closest friends from Germany was very important during last twelve months. Although we could not meet for some time, I never felt lonely or "out of the loop". Skype, interesting short films from Germany, and the good old letter turned out to be very useful means of communication to keep in touch with family and friends, especially from Duisburg, Wulfen, Bonn, Köln and Braunschweig.

This year would have never been such a brilliant experience without the encouragement of my parents. They are my foundation to embrace new challenges and stages in life with curiosity, joy and confidence. Everything I accomplish is related to their trust, support and love.

Table of contents

Abbreviations

AIDS:	acquired immune deficiency syndrome
BYM:	bayesian disease mapping methods
CANZ:	Campylobacteriosis Analysis New Zealand
CAU:	central area units
DHB:	district health board
DPSEEA:	driving force, pressure, state, exposure, effect, action
EHI:	environmental health indicator
ESR:	Institute of Environmental Science and Research
GIS:	geographical information systems
GP:	general practitioner
HIV:	human immunodeficiency virus
IRR:	incidence rate ratio
LCDB:	landcover database New Zealand
LQ:	location coefficient
MAF:	Ministry of Agriculture and Forestry
MAUP:	modifiable area unit problem
mc:	multicollinearity
MoH:	Ministry of Health
MR:	multiple regression
MS:	multiple sclerosis
n/a:	not applicable
NZFSA:	New Zealand Food Safety Authority
NZDep:	New Zealand Deprivation Index
NZPHO:	New Zealand Public Health Observatory
NZWWA:	New Zealand Water & Wastewater Association
NZWERF:	New Zealand Water Environment Research Foundation
OR:	odds ratio
PCA:	principal component analysis
SARS:	severe acute respiratory syndrome
sig:	statistical significance
SNZ:	Statistics New Zealand
std:	standard deviation
SU:	stock unit
TLA:	territorial local authorities
WHO:	World Health Organization
WIS:	weighted index of interseasonal change
WINZ:	Water Information for New Zealand

List of figures

List of tables

1. Introduction

Campylobacteriosis belongs to a group of food and waterborne diseases causing gastroenteritis, which is an inflammation of the gastrointestinal tract. Generally, food and waterborne diseases are associated with a lack of access to clean water and adequate food supplies. This is a particular issue in the developing world because affluent nations have benefited from sanitary, hygienic and technical development to a far greater extent. Table 1 exemplifies several microorganisms causing bacterial, viral or parasitic forms of gastroenteritis.

Table 1: Microorganisms causing gastroenteritis

Pathogen	Disease
Bacteria	
Campylobacter	Campylobacteriosis
Vibrio cholerae	Cholera
Virus	
Rotavirus	Rotaviral diarrhoea
Adenoviridae	Adenoviral diarrhoea
Parasite	
Giardia lamblia	Giardiasis
Cryptosporidium parvum	Cryptosporidiosis

(Kistemann, 1997: 211, modified)

However, within the last few years, highly developed countries have experienced an increase in disease rates related to infections transmitted via food and water. The World Health Organisation (WHO) lists as these reasons for this development (WHO, 2002):

- the globalisation of the food supply,
- changes in microorganisms,
- introduction of pathogens into new geographic areas,
- migration,
- ageing population,
- increase of highly susceptible people and
- changes in lifestyle.

Infection with *Campylobacter* is thought to account for about 5% - 14% of all diarrhoea cases worldwide and is regarded as the most common bacterial cause of gastroenteritis (Mitchell et al., 2002: 1; Weinstein et al., 2000: 42). The spectrum of symptoms includes diarrhoea, abdominal pain, fever, nausea and vomiting (Gatrell,

2002). The epidemiology of campylobacteriosis differs remarkably throughout the world. In tropical developing countries, for example, the illness lacks the seasonal variation observed in affluent nations such as New Zealand, the United States of America (USA) or European countries, and children under two years are especially likely to experience a life-threatening infection (Allos, 2001). In affluent nations, however, severe health effects are rare, and the disease gains importance in terms of its continuous increase of incidence, still widely unexplained in the majority of countries (Coker et al., 2002; Crushell et al., 2004). Since this thesis deals with campylobacteriosis in New Zealand, the focus is on this disease in the context of affluent nations.

1.1 The significance of campylobacteriosis in New Zealand

In New Zealand, campylobacteriosis has been notifiable since 1980. Since then, this disease has been steadily increasing and reached a peak of cases in 2003 (figure 1).

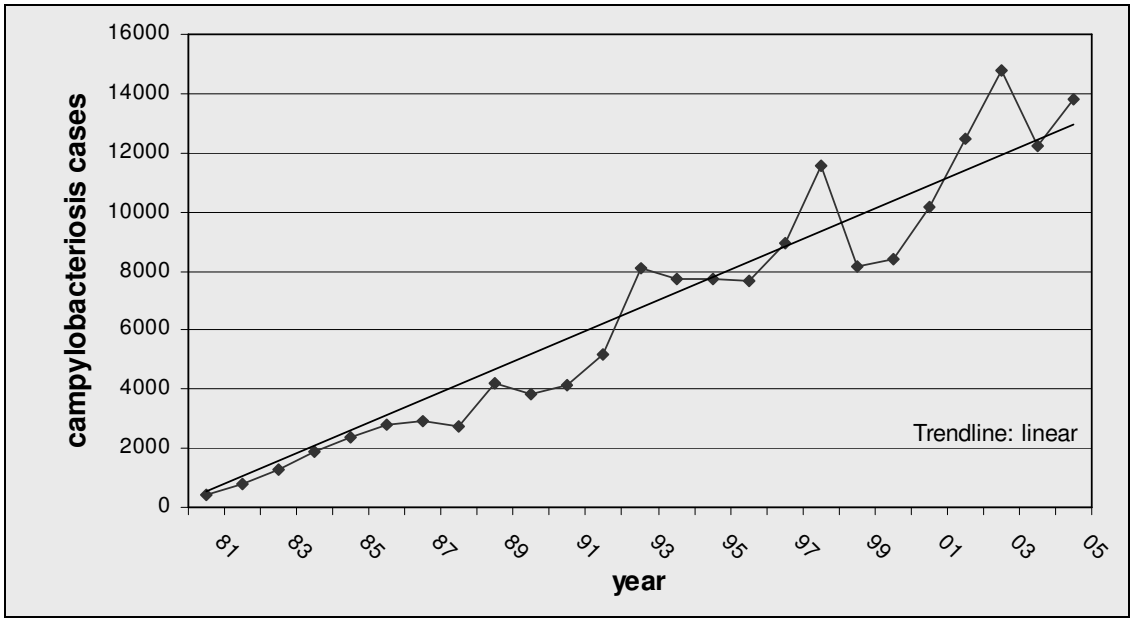


Figure 1: Campylobacteriosis cases in New Zealand
(Brieseman, 1990; ESR, 2006b: 52)

Further, the notifications are unusually high compared to other affluent nations. Figure 2 shows notification rates in selected affluent nations for the year 2003, and it is apparent that New Zealand has by far the highest disease rate (396/100,000). Corresponding rates for 2003 in other countries were: Australia (117/100,000), United Kingdom (UK) (85/100,000), Scotland (87/100,000), Germany (60/100,000) and the USA (13/100,000) (Baker et al., 2006a; Robert Koch-Institute, 2005).

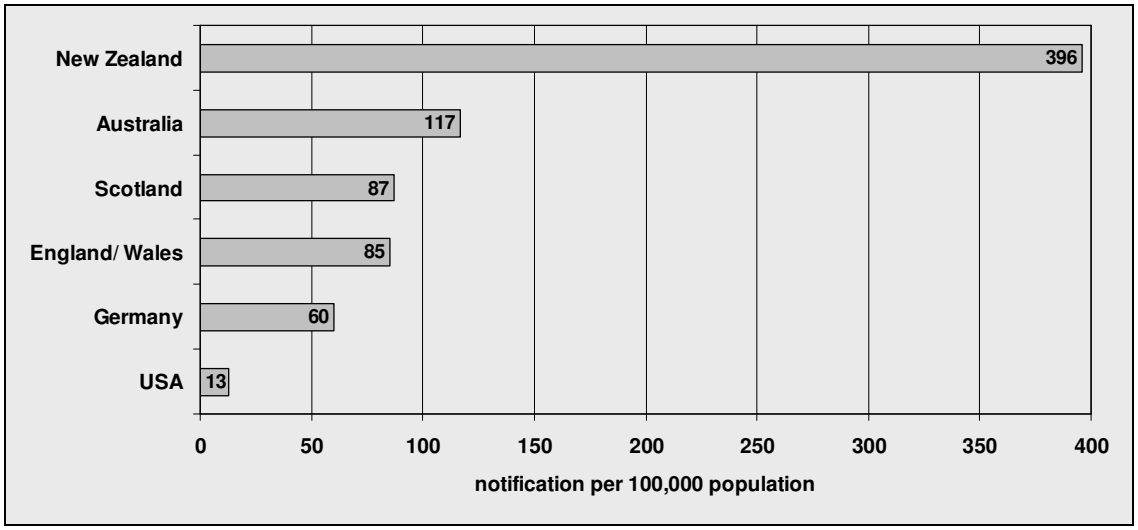


Figure 2: Notification rates for selected affluent nations 2003
(Adapted from: Baker et al., 2006a)

In many countries, such as England, Wales, Australia, and New Zealand, infections with *Campylobacter* now account for more cases per year than salmonellae (Kovats et al., 2005). Although different surveillance systems complicate international comparisons, New Zealand's particularly high rates, which are thought to be the highest reported rates worldwide, still lack a conclusive explanation (Baker et al., 2006a).

Hence, a growing number of people are affected by campylobacteriosis (ESR, 2006b), some suffering from severe impairments to health.[1] Moreover, the economic impact of this disease is significant. The total annual costs from *Campylobacter* alone have been estimated to be about NZ $40,000,000, which equals 73% of New Zealand's total economic cost of foodborne infectious diseases (Scott et al., 2000). Consequently, it is politically and scientifically widely accepted that this persistent situation is neither marginal nor tolerable. However, there is an ongoing debate about the leading causes of this situation and how to deal with it. In particular, the risk of becoming infected from poultry consumption is widely disputed and therefore considered in notably different ways: action alternatives range from improving food-handling to banning fresh poultry from sale.

Current knowledge concerned with the epidemiology and ecology of campylobacteriosis describes a variety of plausible factors contributing to the appearance of campylobacteriosis. The initial cause of developing

[1]The late New Zealand Green Party co-leader, Rod Donald, died of a heart attack at the age of 48 brought on by complications from a *Campylobacter* infection in November 2005.

campylobacteriosis is the invasion of the organism *Campylobacter* into an individual, in which the pathogen reproduces. The circumstances that contribute to the probability of falling ill are multi-causal and vary in time and place. Predisposition, patterns of consumption, the level of exposure and access to health care are examples of different components which are important aspects concerning the susceptibility of the host. Moreover, becoming infected is dependent on the agent's pathogenicity and its virulence (Baker et al., 2006a; Rothman and Greenland, 2005).

Important hosts and transmissions routes of *Campylobacter* are listed below (Devane et al., 2005; TeckLok et al., 2006).

Campylobacter reproduces and survives well in:

- the intestinal tract of animals (e.g. domestic animals, ruminants) and humans,
- food (poultry, raw milk),
- environment (water, soil).

Main transmission routes are:

- consumption of contaminated food or water,
- contact with carriers (infected humans or animals),
- occupational contact with infected meat and other foods.

However, the particular geographical distribution of campylobacteriosis in New Zealand is still widely unexplained. Devane et al. (2005: 981) suggest three main reasons:

> *"Campylobacter transmission routes are complex; studies reported in the scientific literature tend to deal with small aspects of transmission in isolation and rarely involve a cross-disciplinary approach; and, until recently, little research has been conducted into campylobacteriosis in New Zealand."*

Moreover, the WHO (2000) states that one of the major gaps in current knowledge is the quantification of the relative contribution of each identified source to the overall impact of the disease.

With respect to the high morbidity, the traditional public health approach, which generally focuses on the improvement of food handling, water treatment and personal hygiene, seems to be often less than effective (Weinstein et al., 2000). The significant personal suffering and enormous financial costs involved in reducing the risk of becoming infected require further public health and microbiological research to discover more successful preventive strategies (Mitchell et al., 2002).

1.2 Aims and research questions

This study investigates the geographical distribution of campylobacteriosis in New Zealand and the relative importance of factors assumed to be affecting the distribution of this disease, including those related to climate, landuse, water, food, and sociodemographic aspects. The approach aims to explain why certain areas might increase the probability of becoming infected. However, taking into account all aspects affecting the circle of infection is beyond the scope of this thesis. Therefore, it is intended to concentrate this investigation on certain factors of campylobacteriosis that may be of particular interest in the local context. Approaching this topic with a spatial focus will contribute to a better understanding of the causal mechanisms affecting the distribution of campylobacteriosis in New Zealand, a basic requirement for reducing rates and the costs involved.

To meet the aims of the thesis, six research questions are pursued:

Examine the geographical distribution of campylobacteriosis in New Zealand.

- How is campylobacteriosis distributed across New Zealand?
- Is there evidence of disease clustering?
- Where are extreme values of campylobacteriosis located?

Gain insight into the relative importance of plausible determinants assumed to be affecting the distribution of campylobacteriosis in New Zealand.

- Which plausible determinants that might advantage the appearance of campylobacteriosis are important in the local context?
- How are the identified determinants of campylobacteriosis related to the observed *Campylobacter* rates?
- How much of the geographical variation in campylobacteriosis is explained by these factors?

1.3 Structure of the thesis

Chapter two deals with the concepts of environmental health and geographical epidemiology, the theoretical framework of the thesis. Chapter three outlines microbiological key facts and the clinical picture of campylobacteriosis, the development of this disease and its significance for New Zealand. Moreover, conditions affecting the geographical distribution of campylobacteriosis are illustrated in order to establish a background for the methodological approach described in chapter four. There, the analytical approach appropriate to investigate the distribution and the occurrence of extreme values of campylobacteriosis are explained. Furthermore, the application of multiple regression in explanatory research and the analysis of residuals are presented. Chapter five describes the results of the analysis, which are discussed in chapter six. The conclusion summarises key findings of the analysis and puts the results into the context of potential future research.

Figure 3 depicts the different theoretical, methodological and analytical dimensions of the thesis.

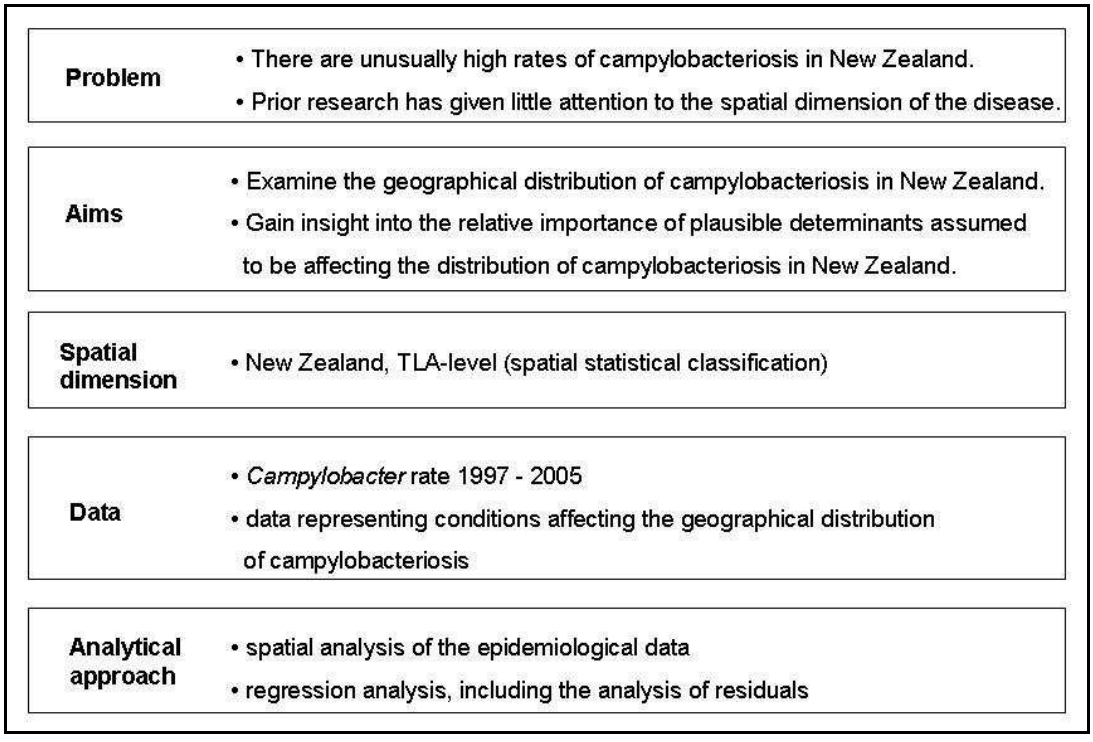


Figure 3: Dimensions of the thesis

2. Principles of environmental health and geographical epidemiology

The concepts of environmental health and geographical epidemiology provide the theoretical framework for the scope of this thesis. It is important to outline the theoretical underpinning for this research in order to explain the chosen methodology and clarify the frame of interpretation for the results of the study.

2.1 Environmental health

The reciprocity of environment and health is described in the medical records of antiquity. Hippocrates (460-370 B.C.) was one of the first who analysed the interrelations between disease, health and environment as he explored the connection between the appearance of waterborne diseases due to different ecological factors. As a result of his studies, he constructed a basic filter in order to use clean water for the treatment of his patients (Gochfeld, 2005: 100; Meade and Earickson, 2002).

According to the WHO (1946: 2) *"health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity".* As widely acknowledged, this definition is rather idealistic and describes a condition that can hardly be achieved in reality. Current definitions attempt to integrate the importance of living a socially and economically productive life. Thus, the maintenance of health requires a permanent adjustment to people's individual associations of wellbeing as well as to the practices of health related governmental decisions (Brown and Duncan, 2000).

Today, the concept of environmental health comprises a broader understanding of "environment". This term integrates not only physical, ecological, socioeconomic and demographic aspects (Mayer, 1996), but also a political component that influences the implementation of health systems, the planning and distribution of health facilities, and our relationship towards sustainable development (e.g. dealing with resources such as water or air). In terms of health policy, responsibility has to be taken for dealing with several health issues due to global change. For instance, the emerging and re-emerging of infectious diseases (SARS, bird flu, foot and mouth disease, hantavirus, water- and food-borne diseases) or the consequences of environmental catastrophes (floods, earthquakes) affect people around the globe. Further, health is determined more and more by socioeconomic factors placed outside the traditional health sector; for example: population growth, formation of

mega cities, migration or the kind of health insurance which is often dependent on one's income (Public Health Advisory Committee, 2004).

How severely these issues influence one's state of health is dependent on both local (internal factors) and global (external factors) environments. On the one hand, predisposition, the body's defences and lifestyles, such as patterns of consumption, are strongly related to our susceptibility to fall ill. On the other hand, political and socioeconomic factors as well as the material wealth of a country play an important role in terms of access to health care. Currently, there is no generally accepted definition of environmental health; instead, there are a variety of meanings that present the diversity of perspectives on environment and health. According to Larry (1998: 2),

> *[e]nvironmental health strives to link environmental quality of both the natural and built environments, with the level of public health and well being. Though lacking consensus in definition, environmental health addresses the interrelationship between human health and the environment. It has been described as the art and science of protecting against environmental factors that may adversely affect human health and environmental quality. Such factors include, but are not limited to air, food, and water contaminants; radiation; toxic chemicals; wastes; disease vectors; safety hazards; and habitat alterations.*[2]

2.1.1 The implementation of environmental health in New Zealand

The implementation of environmental health is mainly based upon the collection and interpretation of health related data obtained by surveillance activities. Surveillance is defined as an *"[…] ongoing systematic collection, analysis and interpretation of […] data for use in the planning, implementation and evaluation of public health practice"* (ESR, 2006b: 3). Some main objectives for disease surveillance include the identification and monitoring of specific disease cases and outbreaks, the investigation of risk factors for diseases, the evaluation of morbidity and mortality trends, and the prediction of emerging hazards (ESR, 2006b).

In New Zealand, surveillance activities are mainly planned and funded by the Ministry of Health (MoH) in collaboration with health professionals in government

[2] This definition is adopted from a list of 28 definitions of environmental health collected by the U.S. Department of Health and Human Services (1998: 1- 6). The quotation combines those factors that are important for the context of this thesis.

departments and scientific institutes. One of the key players concerned with scientific solutions in the area of environmental health is the Institute of Environmental Science & Research (ESR). ESR is a Crown Research Institute, formed in 1992 and owned by the New Zealand government. This institute is in charge of the scientific implementation of environmental health and responsible for a variety of surveillance activities. The three "Public Health Surveillance initiatives and databases" described below are maintained by ESR in collaboration with the MoH and play an important role in the scope of this study.

Public Health Surveillance

Public Health Surveillance (PHS, 2007) provides access to selected public health surveillance data and information. Amongst other functions, it maintains the notifiable disease surveillance database (EpiSurv), which currently covers 50 diseases. Under the Health Act 1956 and the Tuberculosis Act 1948, a doctor is required to report any suspected or diagnosed notifiable disease to the local public health service. There, a medical officer of health classifies each notification either as an individual case or as a potential part of an outbreak[3]. These data are sent to ESR weekly and provide an important basis for health related studies in New Zealand. Furthermore, this link offers information about the New Zealand environmental health indicator project described later.

New Zealand Public Health Observatory

New Zealand Public Health Observatory (NZPHO) is an initiative, developed by the MoH and ESR, to improve access to health data in order to enhance decision-making across the health sector. NZPHO uses Internet and Geographical Information Systems (GIS) technology to make visual population health information widely available. One function of this website is to provide access to the notifiable disease dataset (1997-2005) for different geographical areas, such as district health boards (DHBs) and territorial local authorities (TLAs)[4]. Moreover, it is possible to access several environmental health indicators (EHIs) described in subsection 2.1.2. These data are updated annually.

[3] An outbreak is defined as an epidemic limited to a localised increase in the incidence of a disease, such as in a village, town, or closed institution (Thornley et al., 2002).

[4] DHBs and TLAs are statistical units of New Zealand.

Water information for New Zealand

Water information for New Zealand (WINZ) is New Zealand's drinking water database that provides information on the quality and safety of community drinking water supplies. The data are maintained by ESR and the MoH (ESR, 2006a). For each supply, the register records the name of the community, the sources, treatment plants and distribution zones, the supplier, and the number of people served (figure 4).

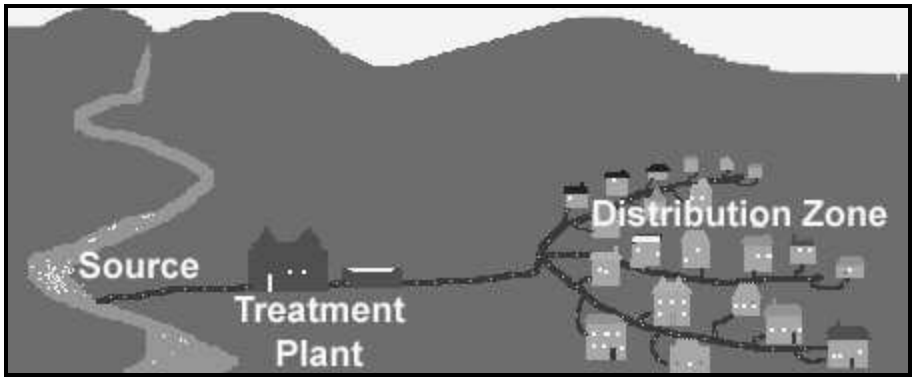


Figure 4: A simple water supply structure
(Source: ESR and MoH, 2006, n.p.)

According to WINZ, the distribution zone is that part of the community receiving water, all of similar quality, supplied by the treatment plant. The source of water is either groundwater (e.g. groundwater well) or surface water (e.g. river, lake). If more than 500 people are connected to one distribution zone, WINZ also records grading for the supply and any substances of public health significance in the supply requiring monitoring (MoH, 2006: 7-11).

Enteric Zoonotic Disease Research Steering Committee

Another important committee, investigating infectious diseases that may be transmitted from wild and domestic animals to humans, is the Enteric Zoonotic Disease Research Steering Committee (2007). This independent working group is closely related to the New Zealand Food and Safety Authority (NZFSA), which is the controlling authority for imports and exports of food and food-related products. The committee was founded as an initiative of the MoH in 2000 to coordinate research on zoonoses across a variety of scientific disciplines and institutes. Its overall aim is the reduction of enteric zoonoses in New Zealand.

In summary, the MoH has the statutory obligation to organise and secure public health. Thus, the ministry is mainly responsible for regulating and funding the political framework concerned with environmental health, whilst several research

institutes provide scientific solutions for its practical implementation. This is dependent on recording notification data and further specific measurements, which are described in the next section.

2.1.2 The measurement of environmental health

In order to measure environmental health, the MoH organises an environmental health project, based on the DPSEEA framework[5] developed by the WHO in collaboration with the European Environmental Agency and the European Commission (1999). This framework consists of six components that describe the interrelation between environment and health (figure 5):

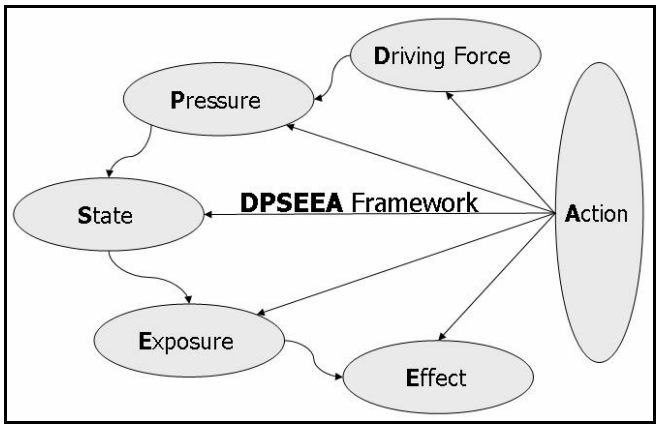


Figure 5: The DPSEEA framework
(Khan et al., 2005: 4)

The (1) *"Driving Forces"* initiate certain (2) *"Pressures"* that impact on the (3) "State" of the environment. In terms of a certain extent of (4) *"Exposure"*, there is a more or less severe (5) *"Effect"* on people's health. In order to avoid or reduce adverse consequences on health, (6) *"Action"* can be taken in each part of the environmental health chain (Khan et al., 2005). Thus, it is assumed that impairments to health can be controlled and reduced through various kinds of activities such as technological innovations or education and awareness raising (Corvalán 1998, in Kremer, 2004).

Surveillance provides the basis for the development of EHIs, which stand for *"an expression of the link between environment and health"* (Briggs et al. 1996, in Khan et al., 2005: 5). Therefore, an indicator represents a variable in combination with a relationship. Measuring this correlation might reveal something about the condition of interest (Briggs and Wills, 1999). Within the New Zealand context, each indicator

[5] DPSEEA: **d**riving force, **p**ressure, **s**tate, **e**xposure, **e**ffect, **a**ction

reflects a particular country specific issue. According to Kahn et al. (2005: 5), there are three types of indicators:

- human health effects that are caused by or associated with environmental exposure (e.g. incidence and prevalence of a disease measured by disease rate);
- measures of environmental quality that have the potential to affect human health (e.g. drinking water quality measured by % of E.coli exceedances in water[6]);
- activities that place pressures on the environment or/ and increase the possibility of exposure in vulnerable populations (e.g. significance of farming across a region measured by proportion of different landuse types).

EHIs facilitate the estimation of environmental health between countries or within statistical boundaries such as DHBs and TLAs. Moreover, these indicators can be used to evaluate the potential benefits of policies. For example, a low intensity of drinking water quality monitoring combined with the occurrence of high incidence rates of water associated diseases obviously argues for investigating the causes. Thus, it is essential to consider the accuracy of a particular EHI in terms of drawing conclusions and coming to decisions: EHIs are not designed to be used as de facto epidemiological evidence. Current ambitions to enhance the indicators' exploratory power are closely related to programmes in Australia and the United States. However, these indicators can be used to generate hypotheses concerning the interdependency between health and environment, one of the major aims of geographical epidemiology (Khan et al., 2005).

[6] *E.coli* exceedance is a general indicator of the state of drinking water quality because its presence in water indicates faecal contamination. If *E.coli* is detected, there is a greater risk of other pathogens being present. Thus, this indicator is a measure of the state of the microbiological quality of drinking water, especially under conditions of inadequate water, hygiene and basic sanitation (Khan et al., 2005).

2.2 Geographical epidemiology

> *Once person, disease, and place are known, we may be able to understand why someone is afflicted and someone else is not. Geographical factors may emerge as paramount in the creation of patterns of distribution (May, 1951: 128).*

In order to demonstrate the special focus of geographical epidemiology[7] within the scope of this thesis, it is first necessary to describe fundamentals of epidemiology important in this context. Subsequently, this section concentrates on particularities in geographic epidemiologic research.

2.2.1 Outlining epidemiology

The term "epidemiology" derives from the Greek: "epi" = upon, "demos" = people/ district, and "logos" = the study of. Thus, epidemiology deals with "science about people" based on the idea that environmental factors influence the occurrence and distribution of diseases (Ahlbohm and Norell, 1991; Centers for Disease Control and Prevention, 2006). So, the primary goal of epidemiology and environmental health is similar. However, environmental health focuses on the concept of health, whereas epidemiology puts the concept of disease at its centre of interest.[8]

2.2.1.1 Research designs and study types

Investigating the human population, studies in early epidemiology were concerned with the aetiology of communicable diseases. This tradition has been maintained till today, and it exemplifies the strong medical focus of epidemiology. Major early epidemiological studies, for instance, contributed to protection against the smallpox virus, which caused millions of deaths until its eradication in 1979. For example, in the 1790s, the English country doctor Edward Jenner vaccinated a healthy boy with a small amount of cowpox material from an infected patient. Six weeks later, he inoculated the child with material from a smallpox pustule, and his "test object" did not become infected. So, Jenner could demonstrate with what would

[7] Within this thesis, "geographical epidemiology", "spatial/ environmental epidemiology" and "disease ecology" are considered as well as synonyms as "geographic correlation studies" and "ecological studies". For further theoretical discussion see Jusatz (1987), Kistemann et al. (1997), and Mayer (1982).

[8] In relation to the WHO-definition of health (WHO, 1946), disease can be defined as an abnormal condition of the body and the mind. For a comprehensive discussion of the concept of disease see Thagard (1997).

nowadays be seen as an ethically questionable research design that a prior cowpox infection conferred protection against the smallpox virus (Beaglehole et al., 2000; Gordis, 2000).

Another achievement of epidemiological studies deals with the investigation of the dramatic increase in the occurrence of lung cancer since the 1930s. The epidemiologist and lung cancer researcher Sir Richard Doll (1912 - 2005) first established a link between smoking and lung cancer in 1950. Subsequent work has confirmed Doll's findings and identified other causes for lung cancer such as asbestos dust and urban air pollution (Beaglehole et al., 2000).

These examples demonstrate some of the variety of study designs applied in the field of epidemiological research. Major epidemiological study designs are summarised in table 2.

Table 2: Major epidemiological research designs

Observational	Experimental
Cross-sectional study	Randomised controlled trial
Case-control study	
Cohort study	

In general, there is a distinction between observational and experimental study designs (Moon and Gould, 2000). Usually, differences in disease rates are observed in a population with the aim of identifying causes for these patterns, and confirming these results statistically or, if possible, experimentally (Schonwalder and Olden, 2003). Experimental or interventional studies, such as the smallpox example outlined earlier, aim to change a disease determinant. Using patients as subjects, this study design can be applied to evaluate new drugs or enhance established treatments for a specific disease (Gordis, 2000). Thus, an interventional study strives to improve its validity by investigating health outcomes according to randomly assigned different exposures (Martin, 2005). However, in order to prove the reciprocity between a feasible pathogenic source and a possibly related disease experimentally, a healthy study group would have to be repeatedly exposed to a putative harmful cause which is ethically questionable and often not applicable. Therefore, many studies follow an observational study design.

Within observational analysis, the researcher measures, but does not intervene. An observational approach is often applied as a first step in an epidemiological study as it generally investigates the occurrence of a disease in a population. For example, Gottlieb et al. (1981 in Beaglehole et al. 2000) described four male patients with a

rare form of pneumonia. The findings were not compared with those of a reference population, but the investigation initiated further epidemiological studies on the condition that became known as acquired immune deficiency syndrome (AIDS). Furthermore, this study design investigates the impact of a number of variables on people's health status (as with the above mentioned investigation of the aetiology of lung cancer).

Summarising research designs in an observational approach, there is a distinction between three major study types (Martin, 2005):

Case-control studies use patients who have fallen ill already and look back to see if there are characteristics of these patients that differ from those who do not have the disease. This study design is suitable to investigate the impact of external (environmental) factors on a specific disease and is thus useful for assessing risk factors.

The cross-sectional study design is applied when there is little prior knowledge about the aetiology of a disease. The researcher questions and examines patients for a specific disease and feasible past and present causal factors.

Cohort studies classify participants who are usually healthy in the first place according to the presence or absence of an exposure. These groups are then followed through time in order to compare the incidence rates between those exposed and non-exposed respectively. This approach requires a long observational study period and is useful for testing hypotheses.

Usually, three sources of bias have to be considered in epidemiologic research (Morgenstern and Thomas, 1993):

- selection bias,
- information bias,
- confounding.

Paraphrasing Morgenstern and Thomas (1993), selection bias occurs due to missing or excluded data (e.g. exclusion of outliers). As a result, the data sample does not reflect the original collected information or the attributes of the population. A major source of information or measurement bias relies on systematic errors during the process of data collection and manipulation (e.g. due to errors recalling data). Confounding refers to a lack of comparability between the exposed and unexposed group. Confounding is a particular issue in nonexperimental research and is also

described as a *"mixing of effects"* (Rothman, 1986: 89). A confounding factor must be a predictor for the disease, but does not need to be an actual cause. For example, a common confounder is age. Age itself is a predictor for many diseases (e.g. rising risk of experiencing a stroke with age), but not a causal factor (Morgenstern and Thomas, 1993; Rothman, 1986).

In summary, this listing specifies some major topics investigated within the scope of epidemiology (Rothman and Greenland, 1998):

- the efficacy of oral antidiabetic medication,
- low-level ionising radiation as a cause of leukaemia,
- hormonal drugs in pregnancy and birth defects,
- passive smoking and its impact on health,
- coffee drinking and pancreatic cancer.

2.2.1.2 Causation in epidemiology

The concept of causation in epidemiology is disputed across several scientific disciplines. For example, epidemiologists criticise laboratory scientists for their restrictive view of causation (Beaglehole et al., 2000). In terms of an infectious disease, the (restricted) cause for becoming infected is the infiltration and reproduction of a pathogenic organism in its host: HIV causes AIDS. Hence, the socioeconomic circumstances, which have an enormous impact on the geographical distribution of AIDS (Kalipeni et al., 2004; Shannon et al., 1991) are not considered as causes by themselves.

The understanding of the concept of causality within epidemiologic research is described using the example of a thrombotic stroke which is an acute neurological injury in which the blood supply to a part of the brain is interrupted. The circumstances leading to this kind of stroke are multicausal. For example, predisposition, prior surgery, high blood pressure, overweight or diabetes are conditions which might trigger the stroke. Epidemiological terminology describes this set of components as "sufficient cause" to initiate the disease. Within the prior mentioned restricted view of causation, a sufficient cause would not be considered as a cause at all. In epidemiology, a sufficient cause is different for each individual and might be the combination of many components or just one (Beaglehole et al., 2000; Moon and Gould, 2000; Rothman, 1986). Although the above mentioned sufficient components, or risk factors, provide information about circumstances

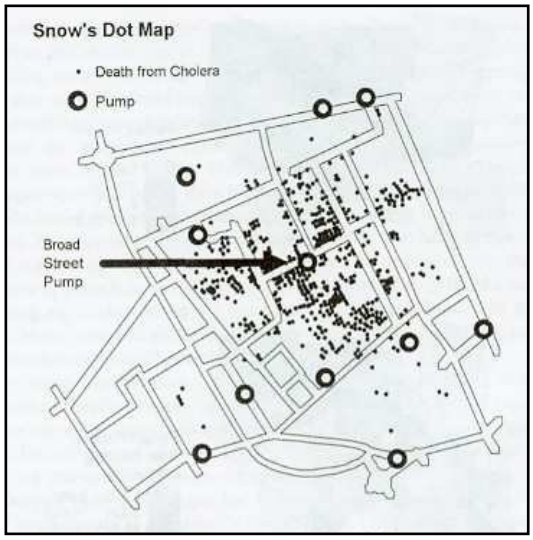
increasing the probability of having a stroke, there is no explanation for the interruption of the blood supply in the first place. This interruption is usually caused by a blood clot which develops due to alterations in the blood flow and its constitution, and injury to the vascular system. Epidemiological terminology describes the existence of the blood clot as the "necessary cause" for a stroke because the disease cannot develop in its absence. From an epidemiological point of view, the mere consideration of the vascular circumstances leading to the development of the blood clot as the cause for a stroke would be insufficient (Beaglehole et al., 2000; Moon and Gould, 2000; Rothman, 1986).

There are various explanations contributing to the understanding of a particular event. A causal factor on its own is often neither necessary nor sufficient (Beaglehole et al., 2000), but the interaction with other factors might cause a disease. Thus, theoretical knowledge of causal relationships provides an understanding of the circumstances, or components in interaction, leading to a particular disease. It is not the explanation of individual causes, but the understanding of their combined interaction that is the basis for prevention and treatment (Manicas, 2006).

2.2.3 Defining geographical epidemiology

"Place" is the theoretical framework for research approaches concerned with the investigation of interactions between predisposing factors, lifestyle and environment (Johnston, 1985; Kearns and Joseph, 1993). Paraphrasing Turnstall et al. (2004), "place" is where a location is, whereas "space" defines what a location is; thus, place is the interpretation of space. This concept is particularly emphasised in geographic epidemiological research, whereas the spatial dimension of disease processes is often neglected in epidemiological study designs focusing mainly on person and time (Rezaeian et al., 2007). In other words, epidemiology generally underlines the medical context of disease, whereas geographical epidemiology stresses their geographical aspects or the "place dimension". However, the methodological approaches, aims and objectives of epidemiology and geographical epidemiology intertwine in various ways (Hertz-Picciotto, 1998). Since geographical epidemiology forms the focus of the thesis, it is proposed to concentrate henceforth on specific aspects of this concept relevant to this analysis.

Techniques of modern geographical epidemiology were first applied by the English doctor John Snow (1813–1858), when he combined epidemiological and geographical methodologies to describe the appearance of diseases. While Londoners were suffering from a cholera epidemic in 1854, Snow considered water as the leading medium for the distribution of the cholera pathogen.[9] Applying the method of disease mapping within a particular London quarter, he identified a particular water pump as the source of infection (figure 6). As people stopped using the contaminated water, the epidemic was stemmed (Mayer, 1982; Meade and Earickson, 2002).



A reconstruction of Dr. John Snow's map of cholera cases in London
(Cromley and McLafferty, 2002: 111)

Figure 6: Cholera map by Dr. John Snow

A very current issue, exemplifying the interrelation between medicine and geography, is the epidemiology of multiple sclerosis (MS). MS is a progressive neurological disease which leads to initially transient weakness, minor motor and visual disturbances, emotional changes, and finally to death. At present, there is no effective treatment for MS except for supportive therapy (Mayer, 1981; van der Mei et al., 2003). Van der Mei et al. (2003) conducted a population based case-control study to examine whether past high sun exposure is associated with a reduced risk of MS. They concluded that higher sun exposure during childhood and early adolescence is related to a reduced risk of the disease, but insufficient ultraviolet radiation may force the progression of MS. The particular spatial aspect results from the fact that this disease has a remarkable geographic distribution. The highest prevalence and mortality rates occur in the temperate zones of both hemispheres,

[9] During this period, air was considered to be the leading cause for the distribution of the cholera agent (= miasmatic theory of cholera).

whilst there are decreasing rates towards the equator. This pattern is associated with geography rather than with race or national origin. It is therefore assumed that climate impacts on the distribution of MS, but the precise interrelation is still unclear (Kurland and Reed, 1988; van der Mei et al., 2003). Hence, MS is a good example of a multifactorial disease because various factors contribute to the distribution of this illness. Comprehensive research is necessary to find an effective treatment. Medical and geographical investigations complement one another as they suggest new findings and solutions on different scales (Mayer, 1981).

Consequently, the relevance of spatial traditions[10] reveals itself very clearly in the area of disease ecology. As clarified above, diseases have geographical patterns; thus, investigating their appearances and non-appearances respectively suggests something about the circumstances which favour or prevent their spread. The crucial concept to explain the emergence of diseases is based upon the ecologic-epidemiological assumption that vector and host coincide at a certain time in a particular place. The factors that contribute to this co-occurrence are multi causal. One of the main questions disease ecology attempts to explore concerns the identification of factors that lead to this spatiotemporal coincidence (Mayer, 1990).

Finally, it needs to be emphasised that diseases show up as individual cases of illness, which can be located in space and time. The identification of localities where diseases occur (geographical aspect) and the explanation of how they are transmitted (epidemiological aspect) are the first steps in revealing their causes and development (Schærström, 1996). Thus, geographical epidemiology deals with the description, explanation and evaluation of geographical patterns concerning disease and death notifications in relation to socioeconomic, environmental and demographic factors in space and time (English, 1992; Meade and Earickson, 2002). Because the approach emphasises medical and geographic aspects, geographical epidemiology is inherently multi disciplinary. Furthermore, depending on the particular research question, this approach also refers to sociology, politics, and natural sciences such as medical geology (Mayer, 1996).

[10] Mayer (1990) refers to Pattison's article (1964) concerning the four traditions of geography and puts these into the context of medical geography. The four traditions are: the spatial tradition, the area study tradition, the man-land tradition, and the earth science tradition.

2.2.4 Methods in geographical epidemiology

Data availability and data structure are determinant factors considering the appropriate methodological study design. Usually, these spatial methods are applied in geographic epidemiological research (Elliott and Wartenberg, 2004; Rezaeian et al., 2007):

- traditional disease mapping,
- cluster detection, clustering and cluster investigation,
- geographic correlation studies.

Elliott and Wartenberg (2004) point out that this categorisation is artificial and does not cover all spatial analytical approaches. Furthermore, this classification overlaps in several ways. Displaying clusters, for instance, results in a final map that is not necessarily solely descriptive, but the visualisation of a prior analytic process. However, this classification is used because it outlines the approaches relevant to this analysis.

2.2.4.1 Disease mapping

Since John Snow's mapping of the addresses of cholera victims related to the locations of water supplies, disease mapping has developed considerably. The main intention of disease maps is the representation of complex data as *"a way of discovering the unexpected"* (Rezaeian et al., 2007: 99). For example, presenting data in maps rather than in tables facilitates the identification of spatial particularities (Elliott and Wartenberg, 2004).

Today, there is a sharp distinction between visualising point data (as Snow did) and area data respectively. On a *point symbol map* (figure 7.1a), a dot represents one or more specific events, such as one person diagnosed with cancer. Accordingly, a concentration of events reveals clustering, or conversely a random distribution. For reasons of confidentiality, disease information is often available only for areas such as districts or countries. Dot density maps (figure 7.1b) use point symbols in order to visualise numbers of events in the corresponding area. Choropleth mapping (figure 7.2) is a statistical approach within disease mapping. These maps depict areas shaded in different colours or patterns to display place-to-place variations (Cromley and McLafferty, 2002). Maps showing the geographic variations in morbidity and mortality of a specific disease are a practical tool to highlight local particularities. This is amongst others useful for initiating or enhancing

surveillance and prevention programmes as the effort and expenses can be concentrated on areas specifically affected (Kulldorff et al., 2006).

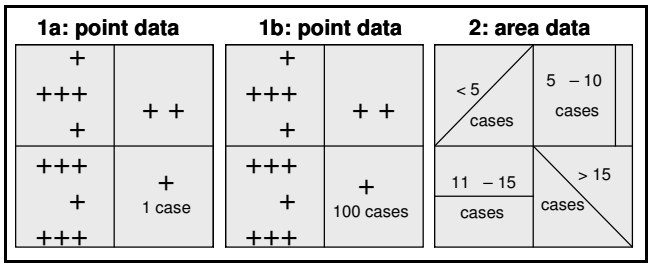


Figure 7: Legends for point and area data

The accuracy of maps is mainly dependent on the availability and quality of the underlying data set. The modifiable area unit problem (MAUP) occurs in association with the choice of projection, size of area units and the number of categories (Openshaw 1984, in Kistemann and Queste, 2004). For example, designing incidence maps is affected by the appearance of extreme values in areas with small populations due to the greater instability of rates. So, the likelihood of extreme values rises within small populations (Hertz-Picciotto, 1998).

Smoothing techniques or the display of P values (= probability maps) are useful to deemphasise such effects (Rezaeian et al., 2007). The aim of smoothing is to reduce variation inherent in the data. If smoothed, peaks and troughs in the data are removed, and the map displays a continuous variation (Lawson and Williams, 2001). In probability mapping, the likelihood of rates is mapped rather than the rates themselves. For example, the location coefficient measures the likelihood of an observed rate in a particular area in comparison to the expected rate of a disease, which is the corresponding national or regional rate (Cromley and McLafferty, 2002).

A recent example of disease mapping is the *Atlas of Cancer Mortality in New Zealand 1994-2000* published by the Ministry of Health (MoH, 2005):

> *The Atlas of Cancer Mortality in New Zealand 1994-2000 shows the recent spatial patterns of cancer mortality in New Zealand. The patterns show areas with […] high or low cancer mortality, but they do not imply the causation of any cancer. No attempt has been made to investigate or suggest possible factors underlying or causing these patterns. However, the maps can be used as a stimulus for further research by suggesting possible aetiological hypotheses (MoH, 2005: 1).*

2.2.4.2 Cluster detection, clustering and cluster investigation

Although traditional disease mapping allows initial conclusions about concentrations of a disease in particular areas, there are specific techniques to identify extreme values or outliers of the disease of interest. Cluster detection is concerned with the exploration of clustering and disease clusters. A variety of definitions is concerned with the terms clustering and clusters respectively (Elliott and Wartenberg, 2004; Morris and Wakefield, 2000; Sabel and Löytönen, 2004; Wakefield et al., 2000). According to Wakefield et al. (2000: 128-129), the term clustering *"refers to the [rather global] pattern […] of disease cases, relative to the pattern of the non-cases"*, while a cluster is defined as a local *"excess of cases"*.

So, the detection of clustering and clusters can be divided into two approaches: global tests provide information about clustering without specifying the location of the clusters, whereas local tests define the location of a specific cluster (table 3):

Table 3: Methods to detect clustering and cluster

Method	kind of test	kind of detection	example
traditional methods	global	detect overdispersion in areally aggregated data	Pearson's chi-squared statistic
distance/ adjacency methods	global	assess the spatial dependence in a set of data	autocorrelation statistics: Moran's *I*, Geary's *c,*
distance/ adjacency methods	local	assess the spatial dependence in a set of data	Getis and Ord's *G** statistic
moving windows	local	detect spatial locations of excess	scan and baysian statistics
risk surface estimations	local	estimate the underlying residual risk surface	kernel estimation

Adapted from Wakefield et al. (2000)

Aamodt et al. (2006) conducted a simulation study of several approaches for detecting disease clusters. They investigated and evaluated amongst others scan (SaTscan) and bayesian disease mapping methods (BYM) for detecting local clusters from a methodological point of view. The authors conclude that the methods differ in their sensitivity detecting clusters, but all methods are generally suitable for identifying clusters with a relative risk larger than 1.5. The practical implementation of scan statistics and kernel estimation is for example applied in two studies estimating the relative risk (Sabel et al., 2000) and investigating spatial clustering (Sabel et al., 2003) of the rare neurological motor neurone disease in Finland. Further research using methods to detect clustering and clusters are concerned with the investigation

of leukaemia (Wray et al., 1999), early outbreak detection (Kulldorff et al., 2005) or the environmental correlates of children's respiratory health (Kingham et al., 1995).

General limits of cluster detection methods are described by Koch and Denike (2001). Paraphrasing the authors, analysing clusters of medical conditions with an unknown aetiology is difficult because the statistical significance of a cluster exists only in relation to further local, social and demographic contexts such as population density and distribution. For example, 10 cases of a particular disease may be normal in a hospital, but constitute an epidemic in a sparsely populated area. Thus, it is crucial to put a potential cluster into its local context. Snow's cholera map, for instance, showed significant clustering at the water pump because the number of cholera cases was noticeably higher there than in neighbouring blocks of houses.

A further approach concerned with the study of clusters is called cluster investigation. Research in this area strives to explain the detected disease clusters, e.g. via regression analysis, linking them to a putative source of hazard (Morris and Wakefield, 2000). Geographic correlation studies are closely related to this methodology.

2.2.4.3 Geographic correlation studies

Investigating plausible aetiological hypotheses is one of the major aims of geographic correlation studies or ecological studies. This study design explores geographic variations of a disease across a defined population in relation to specific environmental health risks (e.g. contaminated water, food), socioeconomic and demographic factors (age, gender, ethnicity), and lifestyle factors (e.g. smoking, diet). The health outcome of these variations is statistically measured on an aggregated geographic scale, such as by districts or countries (Elliott and Wartenberg, 2004).

There are several ways of conducting ecological studies, which are either exploratory or analytic. In exploratory studies, the exposure of interest is not measured, whereas an analytic design quantifies the ecologic association between exposure and disease. A simple exploratory study design compares relative rates across regions and tests for geographic clustering. The analytic approach quantifies an assumed association. Moreover, the variables of interest can be grouped by place (multiple-group study), by time (time-trend study), or by place and time (mixed study). A multiple-group study uses separate sources of data (e.g. census data and

disease rates) to assess the strength of a plausible relationship between the exposure(s) of interest and the disease rate. The time-trend study includes the change in average exposure level over a defined period (weeks, years, seasons), and the mixed study considers both, that is, change over time within groups and differences among groups (Morgenstern, 1995; Morgenstern and Thomas, 1993).

There are several reasons for applying an ecological study design. First, this approach provides the possibility of quickly investigating plausible environmental-health relationships. Even though statistical association is only one part of the argumentation, this procedure has the advantage of avoiding ethically questionable experiments. Furthermore, if the necessary data are available, the execution of these studies uses, in comparison to other research designs, relatively low resources (personnel and monetarily). However, there are issues to be considered when interpreting the results of an ecological study. For example, correlation studies are crucially dependent on scale (MAUP): *"If the regions are large, there is a greater possibility that associations measured at the aggregate level [TLAs] will differ from the same association measured at [a smaller scale such as meshblocks]"* (Rezaeian et al., 2007: 100). Further, the problem of inferring causation from association is known as the ecological fallacy. Issues deriving from the selection bias, information bias and confounding also have to be considered. Moreover, the included variables are often correlated among themselves. Multicollinearity makes it difficult to isolate the true effects of the exposure of interest on the disease (Macfie and Nufrio, 2006).

Hence, the statistical results cannot be specifically related to the individual level as the susceptibility to sicken is closely linked to conditions which never appear evenly distributed within a group (e.g. age, gender, state of health, time of exposure). Nevertheless, a strong statistical association supports the established assumptions. In order to corroborate the hypothesis, further studies on the individual scale are crucial (Morgenstern and Thomas, 1993).

2.3 Summary

The concepts of environmental health and geographical epidemiology were outlined in order to establish the theoretical basis for this thesis. From a meta-theoretical point of view, the DPSEEA framework provides the possibility to operationalise complex environmental-health relationships. This was emphasised by focusing on the implementation and measurement of environmental health in New Zealand.

Geographical epidemiology provides the theoretical and methodological frame to investigate relationships between the distribution of a disease and its potential ecological, demographic and socioeconomic determinants. Different study types were presented highlighting the spatial perspective of geographic epidemiologic research designs. Although ecologic study designs have disadvantages to be considered when interpreting the results, they provide the possibility of analysing environmental health relationships at the aggregated level:

> *"Despite such problems, ecologic studies have been useful in describing differences in populations; even if confounded by unknown or uncontrollable factors, such differences at least signal the presence of effects worthy of further investigation" (Rothman, 1986: 74).*

3. Campylobacteriosis – a persistent and increasing issue

Since this study considers aspects of campylobacteriosis in New Zealand, chapter three focuses on this disease in the context of affluent nations, first considering microbiological, clinical and epidemiological aspects. Subsequently, factors affecting the geographical distribution of campylobacteriosis are discussed, followed by the presentation of a selection of comprehensive risk studies focussing on plausible determinants of campylobacteriosis. The chapter concludes with the subsumption of campylobacteriosis into the DPSEEA framework.

3.1 Microbiological key facts and the clinical picture of campylobacteriosis

In 1886 the German paediatrician Theodore Escherich discovered organisms resembling *Campylobacter* in stool samples of children with diarrhoea. It took another 27 years until the bacterium was clearly identified in foetal tissues of aborted sheep. *Campylobacter* was not successfully isolated until 1972, when microbiologists in Belgium first separated these bacteria from stool samples of patients with diarrhoea (Crushell et al., 2004). Today, sixteen species and six subspecies are assigned to the genus *Campylobacter*. The types most commonly associated with human infections are *C. jejuni* and *C. coli*, causing >80% and >10% of all cases respectively. The name is derived from the Greek “campylos” and “baktron”, meaning “curved rod”, as the organisms are slender, spirally-curved Gram-negative rods (WHO, 2000).

Most *Campylobacters* survive well in the gastrointestinal tract of wild and domestic animals, including animals used for food production. Excreta from infected animals can also contaminate natural water sources and soil (e.g. beach sand). The organism requires specific conditions to grow. *Campylobacter* replicates best in a low oxygen environment at 37°C to 42°C. As the approxim ate core temperature of birds is 41°C to 42°C, they have been considered the prim ary natural hosts of this pathogen. Furthermore, *Campylobacter* occurs in the environment more frequently during the winter months because the organism grows well in water below 10°C. Freezing, drying, acid conditions (pH ≤ 5), UV light and salinity, however, reduce its growth (Altekruse et al., 1999; Newell et al., 2002).

Since the bacteria are part of the normal flora in most animals, they rarely cause illness. Infection of the human body, however, might result in campylobacteriosis diagnosed by faecal culture (TeckLok et al., 2006). Compared to other enteric

diseases, such as giardiasis or cryptosporidiosis, the infectious dose of *Campylobacter* is relatively high. It is estimated that an "optimum" dose for becoming ill is 1,000 - 10,000 cells, whereas <10 - 30 cells alone might cause an infection with *Giardia* or *Cryptosporidium* (Allos, 2001; Percival et al., 2004). Nevertheless, it has to be kept in mind that just one drop of chicken juice may contain 500 infectious organisms (Hood et al., 1988). An outbreak among preschool children involving bird pecked milk bottles also provided evidence that very few organism can cause illness (Mitchell et al., 2002).

Generally, campylobacteriosis is self-limiting and lasts for three to seven days, but children, young (male) adults, the elderly and immune-weakened are specifically at risk of experiencing a more severe reaction to campylobacteriosis (Duncanson et al., 2000; Mitchell et al., 2002). The spectrum of symptoms includes diarrhoea, abdominal pain, fever, nausea and vomiting. Further, people may become carriers while experiencing the disease asymptomatically and are infectious despite not having any symptoms. If treatment becomes necessary, it is mainly supportive with administration of fluid and electrolytes. Antimicrobial treatment is particularly indicated in children with immunodeficiencies and severe cases when there is no response to supportive therapy (Crushell et al., 2004). Usually antibiotics end the symptoms within 72 hours, but growing resistance to these drugs has become a serious worldwide problem (Norström et al., 2006).

While those having experienced campylobacteriosis can obtain immunity to serologically related strains of *Campylobacter*, this disease might also cause severe complications. These comprise infections of the blood, liver or pancreas. Abortion has also been reported as well as the development of Reiter syndrome or Guillain-Barré-syndrome (WHO, 2000).[11] This list represents a selection of diseases caused by *Campylobacter* (Crushell et al., 2004: 8, modified):

- gastrointestinal: enteritis, hepatitis, pancreatitis;
- rheumatologic: Reiter syndrome;
- neurologic: Guillain-Barré-syndrome, meningitis;
- pulmonary: pneumonia;
- intravascular: septic shock, endocarditis, thrombophlebitis.

[11] Reiter syndrome: reactive arthritis (painful inflammation of the joints which can last for several months); Guillain-Barré-syndrome: paralysis that can result in respiratory and severe neurological dysfunction or death in a small, but significant, number of cases (WHO, 2000).

3.2 Campylobacteriosis: New Zealand and the international context

This subsection concentrates on campylobacteriosis in New Zealand embedded into the international context. The focus is on notification trends, outbreaks, sporadic cases, and specific seasonal patterns.

3.2.1 Notification trends: outbreaks and sporadic cases

As mentioned earlier, New Zealand has experienced an increasing prevalence of campylobacteriosis since 1980, and compared to other affluent nations the notification rates are unusually high. For New Zealand, there are 2,568/100,000 cases of campylobacteriosis notified for the period 1997 - 2005. The annual notification rate averages 304/100,000. Figure 8 shows the notification trend of campylobacteriosis for this period peaking in 1998 and 2003.

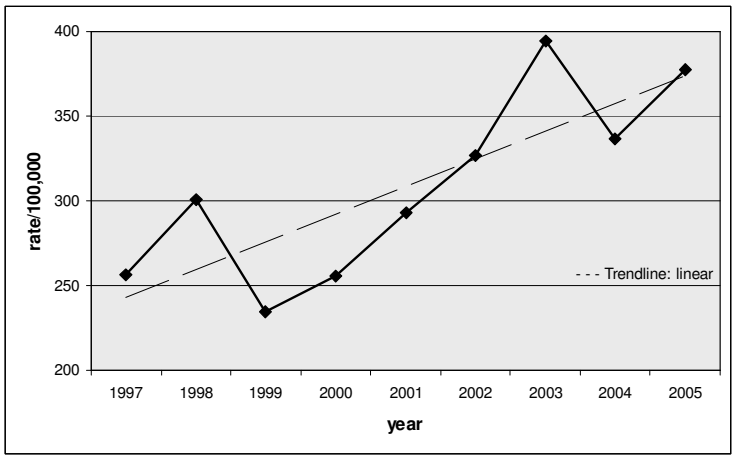


Figure 8: Campylobacteriosis in New Zealand: notification trend

Although a greater level of awareness, improved laboratory procedures and changed surveillance methodologies provide an explanation for the rise of notifications prior to 1990, it is widely accepted that the observed trends of increasing rates are a true effect (Hearnden et al., 2003; McNicholas et al., 1995). Baker et al. (2006a) further conclude that striking similarities between the increase of notifications and hospitalisations additionally demonstrate the reality of this pattern.

There is an important differentiation between two epidemiological patterns of *Campylobacter* infection occurrence: outbreaks and sporadic cases. Usually, outbreaks occur due to the consumption of contaminated food, milk or water, whereas sporadic cases are more likely ascribed to contact with infected pets or animals (Gillespie et al., 2003; Potter et al., 2003). Admittedly, it is often difficult to classify a case definitely into one of these categories, and a number of sporadic cases might be in fact part of an unrecognised outbreak (Gilpin et al., 2006).

Large outbreaks of campylobacteriosis have been reported in several affluent nations, such as the USA, the UK or the Scandinavian countries (Hrudey and Hrudey, 2004). Two of the most severe outbreaks occurred in Washington, USA (1998) and Walkerton, Canada (2000). The outbreak in Washington was traced to well water contaminated by faeces from a septic tank seepage due to heavy rain after a drought. 161 cases were laboratory confirmed, but it was estimated that the total number of affected persons was between 2,800 and 5,000. Two people died of severe complications. People living in Walkerton experienced even worse consequences when already inadequately chlorinated well water was contaminated from cattle manure used on a local farm. 268 cases were confirmed, and more than 2,300 individuals were estimated to have suffered gastrointestinal illness. 65 patients were hospitalised, 27 of them developed severe sequelae, and seven people died (Garg et al., 2006; Hrudey and Hrudey, 2004).

In New Zealand, most outbreaks related to gastrointestinal illness between 2001 and 2005 were registered for the Auckland region, many of them caused by *Campylobacter* (ESR, 2001 - 2005; TeckLok et al., 2006). However, despite the above mentioned possibility of unrecognised outbreaks, campylobacteriosis occurs in New Zealand concordantly with other affluent nations most frequently in the form of individual events (table 4).

Table 4: Outbreak related *Campylobacter* cases in New Zealand[12]

Year	Cases		Outbreaks		
	all cases	cases/100,000 population	number of outbreaks	number of outbreak related cases	outbreak related cases in %
2001	10,148	272	56	301	3.0
2002	12,489	334	50	237	1.9
2003	14,786	396	42	140	0.9
2004	12,213	327	31	130	1.1
2005	13,839	370	47	252	1.8
2001 - 2005	**63,475**	**1,699**	**226**	**1,060**	**1.7**

(ESR, 2001 - 2005)

For the period 2001 to 2005, ESR reported a total of 63,475 cases of campylobacteriosis. Of these, 1,060 cases were related to an outbreak, which approximates 1.7%. With respect to this small percentage and the available data, this thesis does not further consider the investigation of outbreaks and sporadic cases separately.

[12] The annual summary of outbreaks is available since 2001.

3.2.2 Seasonality

The occurrence of campylobacteriosis shows a remarkable annual variation. This pattern, the highest rates occurring in the warmer months and the lowest rates during winter, has been observed in several nations such as New Zealand, several European countries, Canada, and Australia. Nylen et al. (2002) analysed the seasonal variation of nine European countries and New Zealand comparing the time distribution of *Campylobacter* infections based on available surveillance data. The results show that the most prominent seasonal peak occurred in Finland and the least prominent in Scotland and Austria. For New Zealand, the authors describe a less consistent seasonality since the peak was more prolonged.

Kovats et al. (2005) investigated the climate variability and *Campylobacter* infections in 11 European countries, Canada, Australia and New Zealand. They applied regression analysis in order to quantify the associations between the occurrence of seasonal peaks in disease rates in space and time. The results show a distinct seasonality with a peak in spring in *Campylobacter* transmission for most countries. However, the authors could not demonstrate a significant relationship between the rates and their explanatory variables, except for temperature which accounted for 4% of the inter-annual variability of the timing of the peaks. They conclude that climate factors explain some differences in the seasonal variation between countries, and that milder winters might advantage certain transmission routes and the survival of the organism in the environment.

Research concerned with temperature-driven *Campylobacter* seasonality in England and Wales (Louis et al., 2005) examined rates between 1990 and 1999 in relation to weather conditions. The results of the regression analysis showed that temperature, precipitation, and sunshine were significantly related with the *Campylobacter* rates ($R^2 = 0.23$; $p < 0.001$). They also found significant correlations between the rates and various agricultural data (landuse, agricultural labour force, cattle). Surface water, however, did not show a significant correlation with the incidence. Nevertheless, it is worth mentioning that districts with high *Campylobacter* rates were predominantly supplied with surface water which has a higher risk of being contaminated by agricultural run-off or infected animals than groundwater. In contrast, districts with a low incidence were supplied with less than 30% surface water.

The prior investigations demonstrate seasonal tendencies in notification trends in New Zealand being generally consistent with those reported in other affluent nations. Hearnden et al. (2003) also analysed the seasonal variation in notifications of campylobacteriosis in New Zealand, but the authors additionally developed a spatial classification of dominant seasonal patterns for New Zealand's TLAs.

Figure 9 illustrates the seasonal rates of campylobacteriosis in New Zealand for the period 1993 - 2000 per TLA.

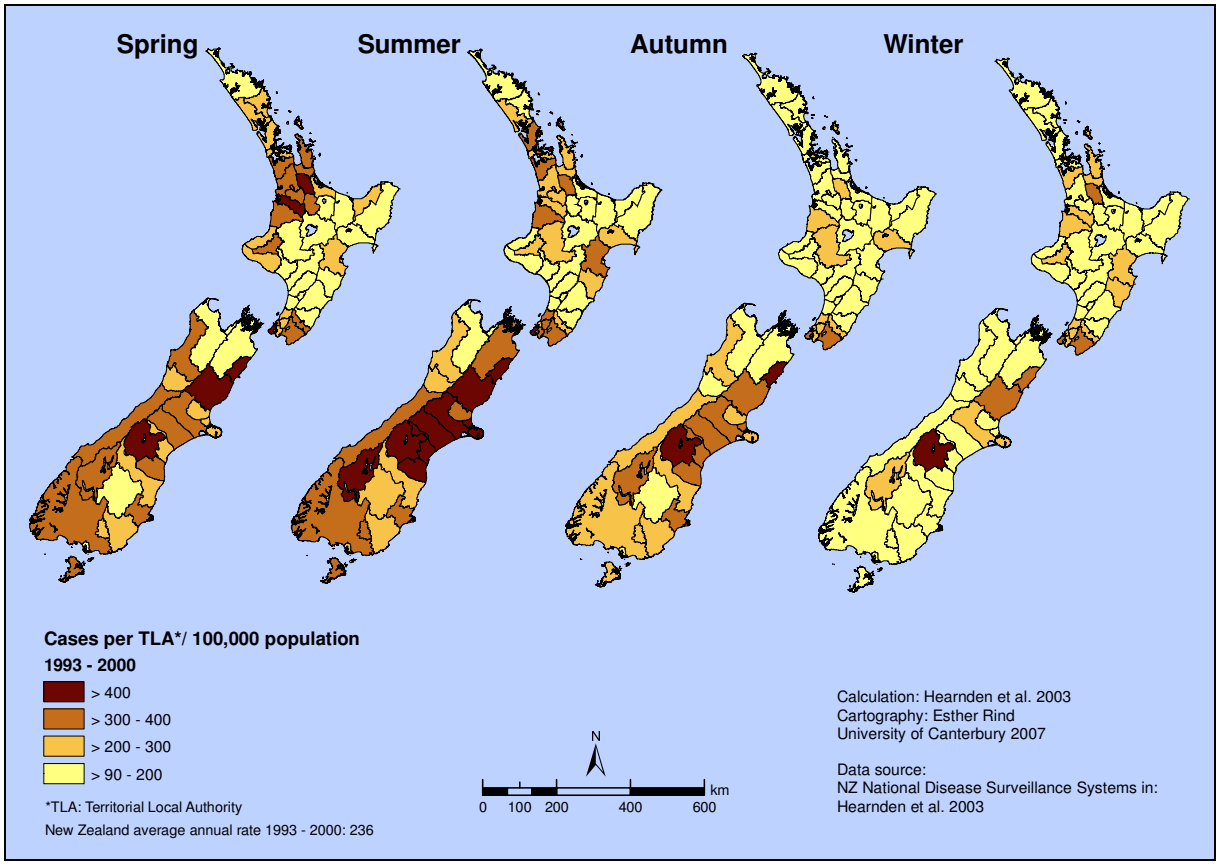


Figure 9: Seasonal rates of campylobacteriosis in New Zealand 1993 - 2000[13]

The authors describe peaking rates in spring and summer. Summer and autumn rates are higher in the South Island than in the North Island, while spring and winter rates show no significant differences between both islands. Moreover, most of the

[13]All rates are calculated per 100,000 population. In order to maintain cartographical clarity there is no TLA-labelling (for TLA-denotation see appendix A1). Immediately prior to the 1989 re-organisation of local government, there were 205 territorial authorities, including cities, boroughs, counties, districts and a town district. These were replaced by 74 territorial authorities, 15 cities and 59 districts, including Chatham Islands County, which retained its county status until 1995 when it became a district known as Chatham Islands Territory. On 1 March 2004, Tauranga District was changed to Tauranga City (SNZ, 2004). As there is hardly any data for the Chatham Islands available, this TLA is excluded from the analysis. If not otherwise indicated, all classification schemes are orientated towards the patterns found using the "natural breaks" or Jenk's optimisation technique. Lighter colours represent lower, darker colours higher values.

South Island TLAs experience decreased rates only during the winter season; whereas on the North Island, lower rates already occur in autumn.

The study further conducts a principal component analysis (PCA) in order to transform the annualised rates for the different seasons into one variable. The result is visualised as seasonal grouping of all rates classified into three different categories representing a particular type of seasonal variation per TLA (figure 10). Three types of summer rates are emphasised in relation to their inter-seasonal variation, and the dominant seasonal patterns are summarised as follows:

- *Low summer incidence and low inter-seasonal variation:* rural North Island, including the South Island TLAs Tasman, Nelson, and Grey.
- *Higher summer incidence and greater degree of seasonality:* urban North Island (including Auckland, Hamilton, Napier and hinterlands, Porirua, Lower Hutt) and the South Island TLAs Marlborough, Tasman, Westland, Waitaki, Central Otago, Clutha, and Southland.
- *Highest summer incidence and greatest inter-seasonal variation:* urban South Island (including Christchurch, Dunedin and hinterlands, Queenstown) and the lower North Island urban centres of Wellington and Upper Hutt.

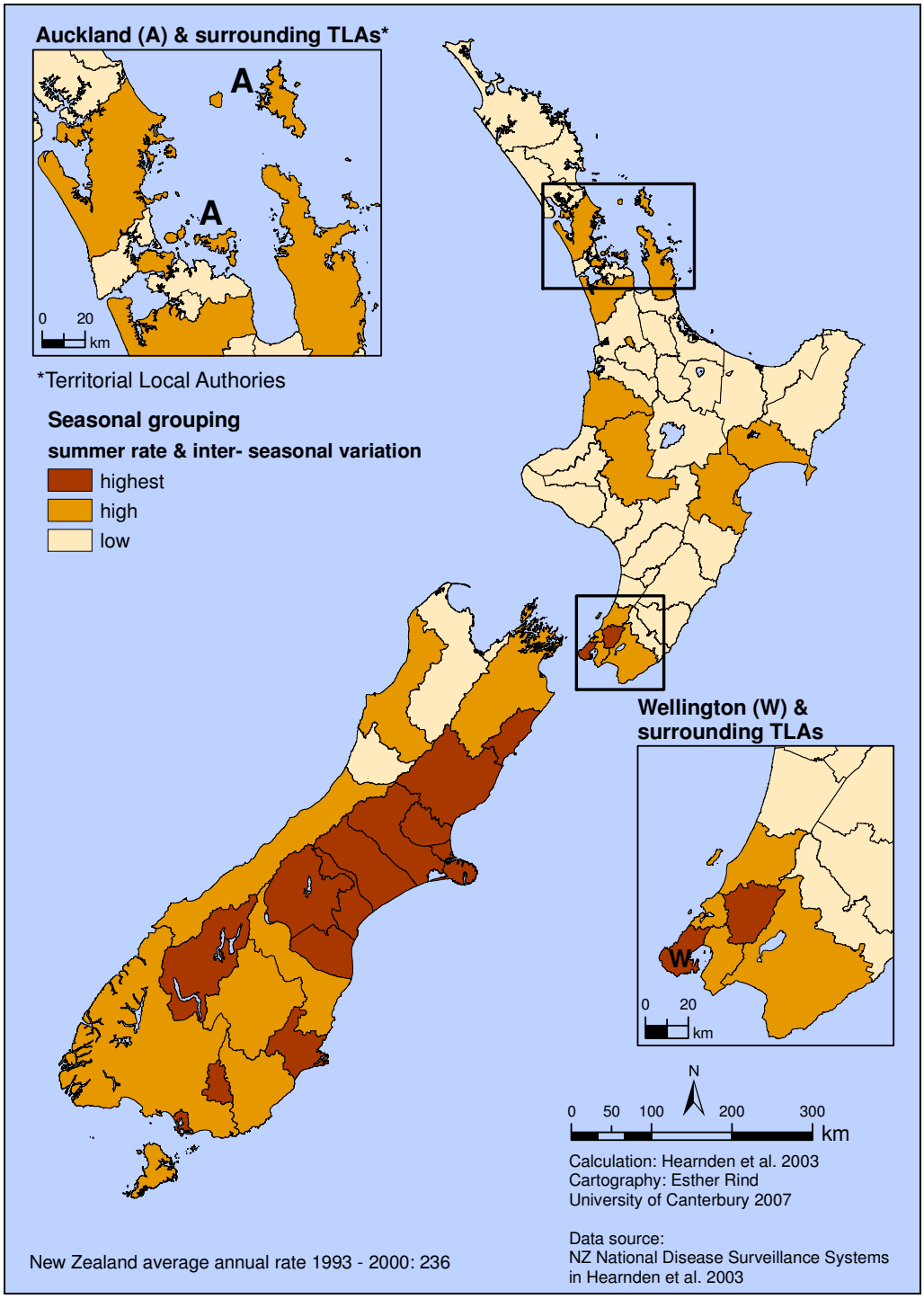


Figure 10: Seasonal grouping of *Campylobacter* rates 1993 - 2000

In conclusion, New Zealand has unusually high rates of campylobacteriosis which vary remarkably with seasons and regions. The notification trends and seasonal patterns are generally consistent with those reported from other affluent nations. The risk factors that might affect *Campylobacter* seasonality are manifold and discussed in conjunction with conditions that generally affect the geographical distribution of campylobacteriosis.

3.3 Conditions affecting the geographical distribution of campylobacteriosis

The conditions that contribute to the distribution of campylobacteriosis are multicausal, and the thesis focuses on certain factors related to this disease that may be of particular interest in the local context. From an epidemiological point of view, these factors can be split into two main groups (Baker, 2006; Beaglehole et al., 2000):

- *underlying epidemiology of campylobacteriosis*, including conditions affecting the distribution of the organism and exposure;
- *surveillance characteristics* contributing to the probability of disease detection, including characteristics of the surveillance system and characteristics of the affected population.

Both of these main categories have subcategories which in turn comprise the conditions particularly of interest for this project (figure 11).

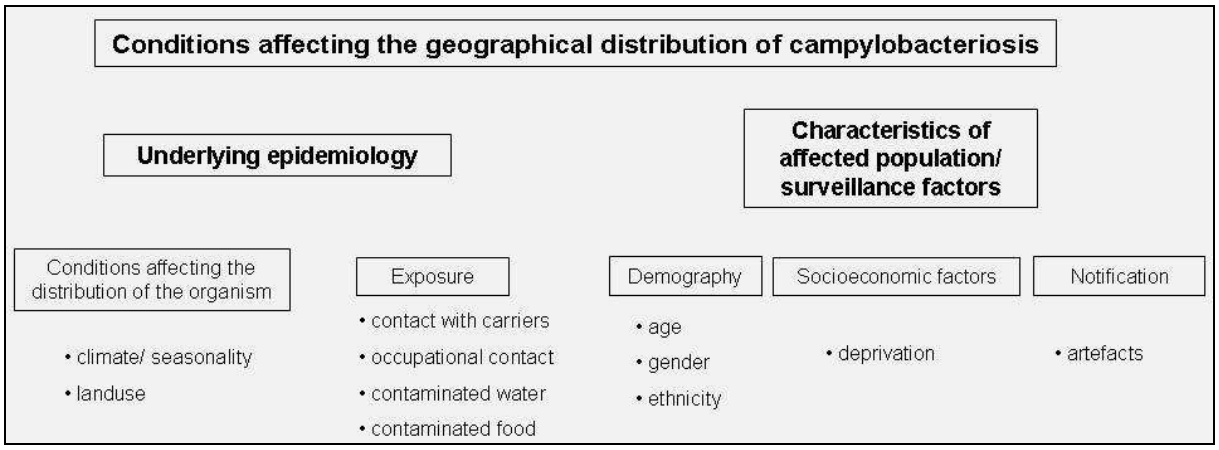


Figure 11: Conditions affecting the geographical distribution of campylobacteriosis

Demographic factors are characteristics of the affected population. In terms of campylobacteriosis they also play an important role concerning susceptibility which is in turn a possible third class in the category “underlying epidemiology”. However, in order to emphasise plausible environmental determinants of campylobacteriosis in New Zealand the above categorisation is chosen. Both of the next subchapters are geared to this overview. To begin with, the conditions affecting the underlying epidemiology of campylobacteriosis are described in detail. Subsequently, conditions related to the characteristics of the affected population and other surveillance factors are explained. This classification provides the theoretical basis for the analysis of this study.

3.3.1 Underlying epidemiology

On the one hand, certain factors such as landuse and climate, including seasonal particularities, determine the more or less distinctive distribution of *Campylobacter* in the environment. In turn, this impacts on the probability of being exposed to a contaminated environment. On the other hand, spatial disparities in exposure to certain health risks such as potentially contaminated food may result in locally varying disease rates.

3.3.1.1 Conditions affecting the distribution of Campylobacter

Climate: rainfall, temperature and seasonal polarities

Rainfall and temperature affect the natural appearance of *Campylobacter* in the environment as these factors codetermine the organism's capability of survival. Heavy rainfall initiates extreme run-off events, and several studies demonstrate the significant increase of microbial loads, including *Campylobacter,* during and shortly after heavy rainfall (Curriero et al., 2001; Hänninen et al., 2003; Kistemann et al., 2002; Louis et al., 2005). Because both low temperatures and low UV radiation advance the pathogen's reproduction in the environment, the survival of the organism is rapidly reduced with rising temperatures during the summer months (ESR, 2001; Hudson et al., 1999; Nygård et al., 2004). Hence, environmental survival of *Campylobacter* interestingly occurs opposite to the peaking appearance of human cases of campylobacteriosis.

However, this typical pattern of *Campylobacter* occurring in the environment more frequently in the colder months is nowadays increasingly discussed as some areas exhibit high pressure in both the colder and the warmer seasons. Eyles et al. (2003) analysed spatial and temporal patterns of *Campylobacter* contamination in the Taieri River (southeastern South Island, New Zealand). Their results show slightly higher concentrations of *Campylobacter* in summer than in winter. They conclude that this oppositional result is due to increased faecal material caused by higher stock densities in summer and seasonal flood events transporting the contaminated load into the river. Additionally, sand appears to be a suitable reservoir for the bacteria during the winter months. Increased run-off and turbulence throughout the summer, e.g. due to increased thunderstorms, can reactivate the pathogen (Pond, 2005). These events might result in a health risk when, for

example, water treatment systems become overloaded and the organism enters drinking water supplies.

Other explanations for the opposite seasonal patterns of the occurrence of *Campylobacter* and campylobacteriosis deal with a change in human lifestyle factors from the colder to the warmer seasons. The summer months come along with more frequent outdoor events such as barbecuing and picnics associated with poorer food hygiene. Further, there is an increased consumption of and contact with potentially contaminated water when hiking, camping and swimming (Gillespie et al., 2003; Nylen et al., 2002) .

Landuse

Climatic factors also codetermine the type of landuse. New Zealand's main land cover class is pastoral (~40%), including tussock and scrub primarily used for extensive grazing (Thompson et al., 2003). This is significant as pastoral farming is a key pressure concerning the distribution of *Campylobacter* in the environment (National Health Committee, 2002). Pastoral farming comprises mainly beef cattle, dairy, sheep and deer farming with sheep farming representing the largest proportion of livestock (77% of the total) (Khan et al., 2005). Livestock production comprises intensive farming (sheep, beef, dairy, deer), particularly on the highly productive lowlands, and extensive high country farming involving just cattle and sheep (Carteris et al., 2000).

Pasture provides about 95% of the diet of dairy cows and nearly 100% of the feed for sheep and beef cattle. Thus, grassland carries a large number of ruminant livestock. Other livestock, such as pigs and chickens are more likely to be penned. This is important concerning people's direct exposure to the organism through the immediate farm environment (Carteris et al., 2000).

The landuse factor especially impacts on water systems as the water quality (e.g. rivers, lakes, groundwater) depends mainly on the type of landuse in the catchment: *"Pastoral agriculture has been implicated as the single largest cause of water pollution in New Zealand"* (Ministry of the Environment 1997 in Khan et al., 2005: 12). Rural run-off increases the number of pathogens in water bodies especially when cattle have direct access to water sources such as rivers or lakes. Murray and Savill (in: Savill et al., 2003) estimate that livestock produce 45 x 10^{12} tonnes of faecal material passing into natural waterways. As mentioned earlier, this is an important issue concerning the existence and effectiveness of the local water

treatment. Thus, the pastoral landuse factor is closely related to different kinds of exposure.

3.3.1.2 Exposure

Water

Water can be a direct causal pathway of becoming infected with campylobacteriosis when people are exposed to contaminated drinking water or recreational water. Water related cases of campylobacteriosis are reported for several affluent nations such as Norway, France, Germany, England, Canada and the USA (Curriero et al., 2001; Evans et al., 2003; Gallay et al., 2006; Garg et al., 2006; Kapperud et al., 2003; Kistemann et al., 2002; Rosef et al., 2001; Sopwith et al., 2003). In particular, Evans et al. (2003) demonstrated a significant health risk from drinking bottled water (OR 1.39; CI 0.98, 1.96)[14]. In terms of recreational water, there is evidence for an increased risk of becoming infected with campylobacteriosis during the summer months when swimming in contaminated lakes or rivers (Pond, 2005). A recent New Zealand study estimated that *"the median proportion of Campylobacteriosis illness that is attributable to freshwater contact recreation is 250/6000, i.e. 1 in 24, about 4%"* (McBride et al., 2002: 23). Moreover, studies from Sweden (Nygård et al., 2004) and the UK (Said et al., 2003) clearly demonstrate that water and environmental factors that influence water quality might have an important impact on the development of campylobacteriosis.

In New Zealand, the overall drinking water standard is very high. In order to provide safe drinking water for the primarily very small drinking water supplies in New Zealand, the Ministry of Health implemented an approach identifying four barriers supposed to protect from water related health hazards (Hrudey and Hrudey, 2004):

- prevention of contaminants entering the raw water of the supply,
- removal of particles from the water,
- inactivation of microorganisms in the water,
- maintenance of the quality of the water during distribution.

However, *Campylobacter* is frequently present in surface waters (e.g. rivers, lakes, dams), and is also detected in groundwater bodies and drinking water. Water

[14] OR: odds ratio, CI: 95% confidence interval

resources are mainly contaminated by agricultural run-off, infected animals having direct access to open water supplies or cross-contamination with sewage due to leakages or insufficient water treatment (Brieseman, 1987; Ikram et al., 1994). This pressure has been increasing because of the recent trend away from sheep farming towards dairying. In addition, increasing urbanisation puts even more pressure on surface and ground water, which are the sources of drinking water (Duncanson et al., 2000; Gilpin et al., 2006).

Savill et al. (2001) found *Campylobacter* prevalent in shallow ground water (75%), river water (60%), roof water (38%), and drinking water (29%). They consider drinking water and recreational water to be a significant transmission route from the environment to humans. In particular, private wells and roof-water supplies in rural areas are at risk of being faecally contaminated from cattle and wild birds respectively as these water supplies remain predominantly untreated (Gallay et al., 2006; Hänninen et al., 2003; National Health Committee, 2002). Examples for water associated cases in New Zealand comprise outbreaks in the township of Ashburton (Brieseman, 1987) and Christchurch (Stehrgreen et al., 1991). In Ashburton, contamination of the town water supply was due to heavy rainfall in conjunction with failure of sufficient water treatment. The outbreak in Christchurch was related to inadequate water treatment as well.

According to New Zealand's drinking water standards (MoH, 2005 -b), *Campylobacter* should be absent from drinking-water samples, and standard disinfection procedures (e.g. chlorination) usually prevent the bacterium's spread in distribution systems. Admittedly, the lack of an endorsed standard method for the detection of *Campylobacter* in water complicates national and international comparisons of laboratory tests analysing samples from different water sources. Conventional tests may underestimate the presence of *Campylobacter* as the bacterium can transform into an undetectable, viable but non-culturable form. For instance, the faecal indicator organism *E.Coli* can be used to estimate the possible presence of *Campylobacter* (Khan et al., 2005; MoH, 2005 -a). However, it has to be considered that *Campylobacter* could still be present in water bodies although *E.Coli* was not detected (Auckenthaler and Huggenberger, 2003; Hrudey and Hrudey, 2004; Rosef et al., 2001).

In addition, it has to be mentioned that besides continuing surveillance improvements, mainly on the part of the Ministry of Health, there still remains a lack

of systematic data, especially concerning environmental factors that provide reservoirs for *Campylobacter* (Baker, 2006). Thus, the true burden of water related infections might be considerably underestimated. This is strikingly proved with the *Annual Review of Drinking-Water Quality in New Zealand 2004* (MoH, 2004) solely using *E.coli* and *Cryptosporidium* as criteria for the assessment of the microbiological health risk related to water. Neither *Campylobacter* nor *Giardia*, both being water related infections in New Zealand, is included in this review.

Contact with carriers

Campylobacter survives well in the intestinal tract of humans and animals. Therefore infected humans or animals are carriers of *Campylobacter*, and through direct contact with carriers there is an increased risk of becoming infected with campylobacteriosis. Although person to person transmission is rather unusual, young adults might be especially at risk when having infected children. Contact with pets or farm animals, however, is a frequently mentioned transmission route (Mitchell et al., 2002).

Campylobacter have been isolated from various birds such as sparrows, pigeons, ducks, and geese (Altekruse et al., 1999; Newell et al., 2002). Intensely farmed poultry is supposed to be frequently infected with the organism. A longitudinal study of *Campylobacter* infections of broiler flocks in Great Britain show *"[…] that more than 40% of flocks were infected with [C]ampylobacter by the time the chicks were 4 weeks old and >90% by 7 weeks"* (Evans and Sayers, 2000: 209). For New Zealand, the information on the prevalence of *Campylobacter* in poultry flocks is extremely limited. Lake et al. (2003) estimate that the pathogen is commonly detected, but compared to other affluent nations the prevalence is suggested to be rather low.

Studies from overseas and New Zealand confirm bovine animals and sheep as being the most important hosts for *Campylobacter* as their intestines are more frequently and more intensely infected with the pathogen. A study from the UK (Stanley and Jones, 2003) examined the natural *Campylobacter* colonisation and transmission among ruminant livestock in the dairy farm environment. They conclude that human infections are significantly related to dairy farm exposure. Similar results are demonstrated by a Danish study (Nielsen et al., 2006). A comprehensive study investigating the appearance of *Campylobacter* in a rural area of New Zealand (Devane et al., 2005; Savill et al., 2003) identified the pathogen in human as well as

in animal faeces, including faeces from dairy cattle, beef cattle, sheep, pigs and ducks. The results indicate that for the examined rural population a range of transmission routes can be related to infected animals, but their relative importance remains to be assessed.

Baker et al. (2006a) review these results and conclude that they might explain a small proportion of New Zealand's high rates of campylobacteriosis. Additionally, they view the findings as being predominantly important for explaining high rates in children living in rural areas as they are supposed to be especially at risk of becoming infected by contact with infected dogs, cats, and farm animals. This in turn provides an explanation for the seasonality of infections which peak in spring together with increased farming activities such as lambing and calving (Baker et al., 2002; Mitchell et al., 2002). The latter is also discussed in conjunction with fly transmission of *Campylobacter*. The seasonal increase of campylobacteriosis in England and Wales is supposed to be partly caused by direct or indirect human exposure to material carried by flies that have been in contact with infected faeces or contaminated foods (Nichols, 2005). The importance of this transmission route for New Zealand is disputed, but has not been quantified in the New Zealand context (Campbell et al., 2006; Nelson and Harris, 2006; Wilson, 2005).

Occupational contact

Closely related to the prior argument is the risk of being infected by occupational contact. People such as farmers, farm and meat processing workers, abattoir and meat freezing plant workers, and veterinarians are particularly at risk of becoming infected because their work brings them into close contact with potentially infected animals (Kapperud et al., 2003; TeckLok et al., 2006).

Several studies found more or less significant evidence for occupational contact being a particular risk of contracting campylobacteriosis. Results from an English study of primary indigenous sporadic cases of *Campylobacter* infection showed that occupational exposure to raw meat is a significant independent risk factor for falling ill (OR 9.37; CI 2.03, 43.3) (Adak et al., 1995). A prospective case-control study investigating risk factors for sporadic *C.jejuni* infections in rural Michigan concludes that persons engaged in poultry husbandry had particularly increased odds of campylobacteriosis (OR 6.9; CI 1.4, 33.0). They further suggest that about 18% of cases occurring in rural populations are attributable to poultry husbandry (Potter et al., 2003).

In New Zealand, Brieseman (1990) refers to earlier studies suggesting that meat handlers, the unemployed and housewives were the occupational groups in which the highest numbers of cases were found. Further research did not confirm these findings though. This is in accordance with a study from England (Adak et al., 1995) demonstrating that occupational contact with livestock or their faeces was a protective factor (OR 0.44; CI 0.21, 0.92). Eberhart-Phillips et al. (1997), however, verify a statistical association between occupational contact and bovine carcasses (OR 2.83; CI 1.12, 7.19).

Food

Food is supposed to be the most significant factor in the circle of infection of campylobacteriosis. Most studies focus on the prevalence in poultry, where the proportion of contaminated products is known to be very high (Keener et al., 2004). However, other foods such vegetables, fruits, red meat, seafood, garlic butter, and raw milk were found to be contaminated as well. Generally, it is observed that the survival of *Campylobacter* in food is better the colder it is stored unless it is frozen (Baker et al., 2002; Reiersen et al., 2002).

In New Zealand, the prevalence of *Campylobacter* in raw and ready-to-eat chicken products usually ranges between 50-70% (Baker et al., 2002; Lake et al., 2003). A study conducted by ESR investigated 300 retail packs of fresh poultry from supermarkets in Christchurch. Of the 300 packs sampled, 32 were positive and 24% were externally contaminated. It is concluded that cross-contamination in supermarkets may play a significant role in the distribution of campylobacteriosis (ESR 2002 in: Lake et al., 2003). A further study from New Zealand also tested raw chicken from retail outlets and raw milk samples (Hudson et al., 1999). Of the 113 chicken samples, over 56% were positively tested for *Campylobacter*, but only one milk sample proved to be contaminated. The authors found their results in agreement with a study conducted in New Zealand by Campbell and Gilbert (1995). Additionally, Ikram et al. (1994) found that when chicken was eaten at home this turned out to be a protective factor (OR 0.36; CI 0.14, 0.9), whereas consuming chicken at a barbecue was a risk (OR 3.00; CI 0.99, 9.34). The authors conclude that undercooked or poorly handled chicken is a significant source of human infection with campylobacteriosis. Moreover, eating chicken at a restaurant (OR 3.85, CI 2.52, 5.88) or from a takeaway establishment (OR 1.70, CI 1.24, 2.32) is also considered to be a significant risk for becoming infected (Eberhart-Phillips et al., 1997).

Other food associations with campylobacteriosis are investigated in a study from Christchurch. Brieseman (1990) showed that the highest number of infected people had consumed chicken (58%), followed by seafood (33%), café sandwiches (29%), meat pies (26%), and raw milk (6%). This is in accordance with international study results from other affluent nations such as the USA (Friedman et al., 2004), Canada (Michaud et al., 2004), England and Wales (Adak et al., 2005; Frost et al., 2002; Gillespie et al., 2003; Sopwith et al., 2003), Finland (Schildt et al., 2006), and Norway (Kapperud et al., 2003). In summary, these foods associated risk factors are identified in the mentioned studies:

- preparing any kind of raw meat in the kitchen,
- consumption of chicken and non-poultry meat in a restaurant or takeaway,
- eating any kind of meat or fish at barbecues,
- eating raw vegetables, fruits, and berries,
- drinking raw milk or raw milk products,

3.3.2 Characteristics of affected population and surveillance factors

As subsequently demonstrated, several surveillance factors potentially influence the probability of disease detection and consequently the notification data of campylobacteriosis.

3.3.2.1 Demography

As in other affluent nations, the average annual incidence in New Zealand is significantly higher in males (326/100,000) than in females (296/100,000), particularly in children aged 0 to 4 years and young adults aged 20 to 29. The rate of infection appears to decline conclusively in risk after the age of 45 years (Butzler, 2004; Peterson, 1994). Europeans show the highest notification rates (329/100,000) compared to Maori (94/100,000), Pacific People (70/100,000) and people of "other ethnicities" (193/100,000) (Baker et al., 2006a; ESR, 2002; Mitchell et al., 2002)[15].

The reasons for children and young males being particularly susceptible to developing campylobacteriosis have not been clarified. Explanations range from underlying immunological differences between the genders, and children being more likely to see a doctor and thus their infection being notified more frequently.

[15] All figures were calculated by Baker et al. (2006a) for the years 2001 - 2003, and they reflect a continuous trend for the overall notification period of campylobacteriosis

Moreover, men are supposed to face greater exposure through occupational contact, less stringent kitchen hygiene, and transmission via homosexual contact (Allos, 2001; Tauxe et al., 1988; WHO, 2000). Ethnic differences are frequently linked to socioeconomic deprivation, which is discussed in the next subsection. Because ethnic differences are less distinct for campylobacteriosis hospitalisations, Baker et al. (2006a) suggest that some of this observed difference is related to poorer access to primary care and health services resulting in lower rates of notified cases.

3.3.2.2 The impact of socioeconomic factors and surveillance artefacts

Socioeconomic factors are considered to have a significant influence on health and ill health respectively. In terms of campylobacteriosis only a few research projects link this disease directly to social deprivation. Both of the subsequently mentioned studies relate social deprivation particularly to the higher risk of becoming infected through contaminated water supplies. The National Health Committee of New Zealand (2002) considers people living in small rural communities to be less likely to have access to safe, regularly monitored water supplies than those in urban areas. They further conclude that the most deprived communities tend to be at highest risk from unsafe water supplies, and Maori are over-represented in these communities. Thus, small rural communities are more likely to experience higher water related disease rates, specifically from failed or inadequate sewage disposal systems.

Having investigated the public health risks associated with community water supplies in relation to social status in New Zealand, Hales et al. (2003) also conclude that deprived communities in New Zealand might experience a disproportionate burden of adverse health effects as a result of poor water quality. However, their results show a particular risk for urban areas, where the odds of water supplies being high risk were 3.76 times greater for the most deprived decile compared with the least deprived decile (95% CI: 2.95 to 4.78).

In conclusion, empirical findings dealing with the relation between disease, health and socioeconomic disadvantage are frequently inconsistent (Blakely et al., 2006). Moreover, there is little evidence showing a distinct association between social deprivation and campylobacteriosis. This requires further investigation.

As mentioned earlier, Public Health Surveillance maintains the notifiable disease surveillance database of New Zealand, and the notifications of campylobacteriosis

by medical practitioners have increased in the last twenty years. Studies have confirmed that this overall effect cannot be described as a surveillance artefact (Baker et al., 2006a; McNicholas et al., 1995). However, from a spatial point of view it is not unlikely that the occurrence of local disparities in disease rates might be due to various surveillance artefacts.

In terms of the organism's detection, different laboratory or reporting procedures might influence the number of notifications. Potential measurement errors might cause a false positive and false negative confirmation of cases respectively. Further, different demographic factors and social deprivation might favour the probability of certain population groups being notified or remaining unnotified (Sneyd and Baker, 2003; Wilson, 2005). A study investigating infectious intestinal diseases in England (Wheeler et al., 1999) concludes that only one out of six people who become infected present to a general practitioner (GP). Amongst others, this might be associated with the accessibility to health care institutions expressed, for example, in GPs per population, most notably decreasing in remote areas. The authors further argue that the proportion of cases not notified is large and varies enormously by the particular organism and the affected population.

These arguments exemplify inherent limits of surveillance data. This is an issue inasmuch as surveillance artefacts might be responsible for a significant part of regional variation in disease rates, but this is exceedingly difficult to quantify. Therefore, notification data might neither reflect the true burden nor the real geographical distribution of a particular disease (Sneyd and Baker, 2003; Wilson, 2005). Being aware of these limitations, it is nevertheless possible to either confirm already validated results or to reveal a prior unconsidered relationship that has to be further investigated using different research methodologies.

3.4 Comprehensive risk studies with a particular spatial focus

Research approaches investigating a wide spectrum of transmission routes combined with a particular spatial approach are rare as it is extremely difficult to quantify the relative contribution of each plausible determinant of campylobacteriosis to the overall prevalence of the disease. Nevertheless, there are international and national examples applying a particular spatial research design.

3.4.1 International focus

A study from Sweden (Nygård et al., 2004) investigated associations between the geographical distribution of *Campylobacter* incidence and environmental factors related to water and livestock. Regression analysis was used to estimate the strength of the associations. The univariate regression analysis significantly relates these variables to increased *Campylobacter* incidence: living in a rural area; ruminant, swine and poultry density; water-pipe length (m/person) and mean annual temperature. The proportion having access to public water supply was inversely related to the rate. Multivariate analysis identifies ruminant density (IRR 1.08; CI 1.05, 1.11)[16], water-pipe length (IRR 1.12; CI 1.08, 1.16), and mean annual temperature (IRR 1.05; CI 1.03, 1.07) as independent risks for an increased *Campylobacter* rate. The results indicate that water might be an important transmission route for *Campylobacter* infections in Sweden, especially through potentially contaminated water distributed by public water supplies. The authors conclude that the findings need to be confirmed by further investigation considering in-depth microbiological and epidemiological aspects.

A recent study from Scandinavia examines Norway's incidence trend and risk factors for *Campylobacter* infections for the period 1999 - 2005 (Sandberg et al., 2006). The focus is on estimating the risk for food, water, contact with animals and climatic factors (rainfall and temperature). The final regression model, however, only includes treated drinking water as a protective factor (IRR 0.776; CI 0.630, 0.954) and rainfall as a risk factor for becoming infected (IRR 1.006; CI 1.005, 1.007). Because the model did not reveal any geographical clustering of cases, the authors conclude that there are not any particular local risk factors for becoming infected. They evaluate their results as being consistent with the assumption *"[…] that several risk factors are of differing importance at different time for the incidence of sporadic Campylobacter infections"* (Sandberg et al., 2006: 8).

Morgan (2002) undertakes aspatial and spatial analysis in order to explore *Campylobacter* cases notified in South Australia between 1990 and 1999. The aspatial part of the investigation identified two age groups and described the temporal variation in rates for the whole of Australia. Notifications for children were significantly higher than for adults. Gender differences are confirmed only for

[16] IRR: incidence rate ratio

children below 15 years. The subsequently conducted surface analysis and the space-time cluster analysis show that the timing of infection may be closely related to environmental variables affecting the survival and distribution of the organism. The author concludes that the seasonality of infections is controlled by environmental factors. The spatial distribution, however, is mainly dependent on socio-economic factors related to surveillance factors and not the illness itself.

3.4.2 New Zealand focus

Explaining differences in local disparities of campylobacteriosis with climate, agriculture and water catchment as explanatory variables is described by Hearnden et al. (1998). The usefulness of using GIS to examine the spatiotemporal distribution of campylobacteriosis in relation to various ecological factors is investigated by Eyles (1999). The author analysed notification data of campylobacteriosis for the period 1993 to 1998 on the TLA-level. The results show three distinct seasonal patterns of campylobacteriosis:

- low incidence rates both during the summer and the winter periods in Northland and Central North Island TLAs;
- high incidence rates in the summer months only in Christchurch City;

It is concluded that GIS has been used primarily as a hypothesis-generating tool.

Mitchell (2000) used GIS to determine whether there is an association between water quality and enteric diseases in Canterbury. The author compared rates of notified enteric diseases from different water distribution zones in Canterbury to investigate a potential relationship between water quality and the disease rates. The results show a correlation between the enteric disease rate and water quality ($R^2 = 0.89$). In terms of data accuracy, the author mentions these concerns:

- the WINZ database was often incomplete,
- some TLAs did not have accurate population data for the water distribution zones in their districts,
- the shape files of the zones were not always accurate,
- the grading of some supplies had not been reviewed for up to ten years.

Mitchell concludes that despite these limitations the results are consistent with previous research in the North Island also confirming the significant role of contaminated water supplies in notifications of enteric diseases.

3.5 Summary

It has been shown that campylobacteriosis is a continuous and increasing issue in New Zealand. Various factors which are particularly important in the New Zealand context of campylobacteriosis were presented considering the international context as well. Comprehensive international and national risk studies identified and assessed several important transmission routes for *Campylobacter*. Summarising the prior arguments and exemplifying ongoing trends, these transmission routes were recorded for New Zealand for the year 2003 (Morrison and Smith, 2004):

- foodborne: 56% (2303/4084),
- contact with farm animals: 27% (1283/3514),
- untreated water: 19% (766/4077),
- recreational water contact: 16% (711/4533),
- overseas travel mainly in combination with the consumption of contaminated food and water or recreational water contact: 7% (345/5298).

Additionally, one of the most comprehensive case-control studies also identified persistent demographic and socioeconomic characteristics for campylobacteriosis cases in New Zealand (Eberhart-Phillips et al., 1997):

- ethnicity: European (>90% of all cases),
- household income: >40,000 NZ $/y (>50% of all cases),
- highest qualification: university degree or technical qualification (>40% of all cases),

Further, the same study assesses these odds ratios (OR) for important food and non-food exposures (examples):

- consumption of poultry: OR 1.31 (CI 1.03, 1.67);
- rainwater source for home water supply: OR 2.20 (CI 1.04, 4.65);
- handling of bovine faeces: OR 3.09 (CI 1.57, 6.10);
- pets at home with diarrhoea: OR 2.08 (CI 1.05, 4.15);
- contact with cattle: OR 2.29 (CI 1.03, 4.05);
- occupational contact with cattle carcasses: OR 2.83 (CI 1.12, 7.19).

These findings are in accordance with results from comprehensive risk studies from England, Wales, Norway, the USA, and Canada (Gillespie et al., 2003; Kapperud et al., 2003; Michaud et al., 2004; Potter et al., 2003; Sopwith et al., 2003; Stanley and Jones, 2003).

As already outlined, the DPSEEA framework provides a possibility to operationalise complex environmental-health-relationships such as the epidemiology and ecology of campylobacteriosis (figure 12).

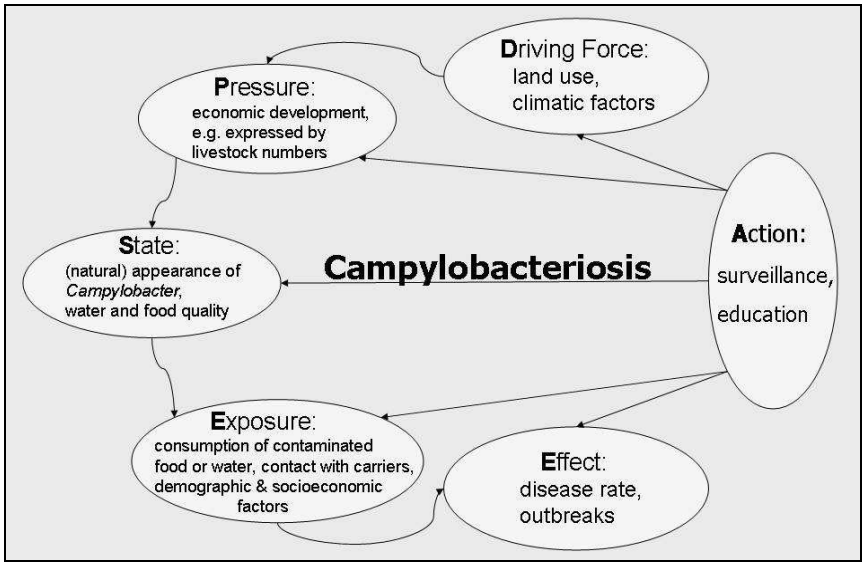


Figure 12: Campylobacteriosis within the DPSEEA framework
(Khan et al., 2005: 4, 9, modified)

Emphasising for instance the water related perspective of campylobacteriosis, one of the *driving forces* is the type of landuse combined with the impact of climatic factors, such as heavy rainfall. With the *pressure* of growing livestock numbers, the *state* of the water quality is increasingly jeopardised when farm or domestic animals have direct access to catchments of open water supplies. People's *exposure* to consuming contaminated water is, for example, determined by the kind and extent of water treatment. These interrelated factors have a direct *effect* on the extent of the disease rate and number of outbreaks. *Action* to deal with campylobacteriosis comprises mainly surveillance and educational programmes.

The prior arguments clearly demonstrate the need for a comprehensive study approach in order to consider the complex transmission routes of *Campylobacter* and to estimate the relative contribution of plausible factors that either advantage the appearance of the organism in the environment or increase the probability of becoming infected with campylobacteriosis.

4. Methodological approach

Recent geographical epidemiologic study designs aim to combine the traditional approaches of disease mapping, cluster detection and ecological analysis into a comprehensive spatial approach to research feasible cause and effect relationships (Elliott and Wartenberg, 2004; Morgenstern and Thomas, 1993). Integrative analysis in epidemiologic geographical research is mainly due to advances in GIS which provide comprehensive possibilities to explore, manipulate, analyse and visualise data from various sources (Clarke et al., 1996; Wall and Devine, 2000). The analysis of this study consists of five steps (Elliott et al., 1995; van den Berg and von der Ahé, 1997):

- exploring evidence of disease clustering,
- detecting areas of disease excess,
- investigating plausible causes,
- testing plausible relationships,
- comparing, presenting and evaluating the results.

The research is conducted on the TLA-level, as the study investigates the geographical distribution of campylobacteriosis for the whole of New Zealand, and the necessary data available for the scope of this thesis are obtainable on this spatial scale.

4.1 The analysis of the epidemiological data

The first aim of the thesis, **"Examine the geographical distribution of campylobacteriosis in New Zealand"**, requires a methodology appropriate to describe the distribution of campylobacteriosis and to identify extreme values of this disease for the study period 1997 - 2005 on the TLA level. Hence, the analysis of the epidemiological data comprises four steps:

- visualisation of the crude *Campylobacter* rate,
- visualisation of the change of incidence in campylobacteriosis,
- investigation of disease clustering,
- identification of outliers.

4.1.1 The distribution of campylobacteriosis across New Zealand

The visualisation of the crude *Campylobacter* rate and the change of incidence in campylobacteriosis for the period 1997 - 2005 are calculated and displayed in order to meet the first research question of the thesis: *How is campylobacteriosis distributed across New Zealand?* The campylobacteriosis notification data (cases per TLA) for the years 1997 - 2005 were received from NZPHO and added together in order to obtain a greater population.[17] These data are visualised as annual crude rate per 100,000 people. Due to data restrictions, demographically standardised rates cannot be calculated for this project. Hearnden et al. (2003) tested the impact of age standardised rates for campylobacteriosis over New Zealand's TLAs. As they found little difference between crude and standardised rates, they decided to use the crude rates for their analysis.

The results are displayed as choropleth maps. As mentioned earlier, if not otherwise indicated, all classification schemes are orientated towards the patterns found using the Jenk's optimisation technique. Here, classes are based on natural groupings inherent in the data. Break points are identified by choosing the class breaks that best group similar values. Thus, the classes are set where there are relatively big jumps in the data values. Generally, lighter colours are used to represent lower values, whereas darker colours indicate higher values. A red-green colour scheme is used to indicate increase or decrease of values (Bartels and Van Beurden, 1998).

Moreover, the Pearson correlation coefficient (r) is calculated in order to investigate whether there is a potential statistical association between the extent of the *Campylobacter* rates and their development over time (1997 - 2005). The Pearson correlation coefficient is calculated using the actual data values and ranges from -1 to +1. Values close to -1 indicate a strong negative linear association between the variables, whereas values close to +1 indicate the opposite effect. If the result of the correlation is close to 0, there is no statistical association between both of the variables. A p-value below 0.05 indicates statistical significance of the association.

[17] If not otherwise indicated, the calculations were conducted using Microsoft® Office Excel and Access 2003.

4.1.2 Evidence of disease clustering and identification of outliers

As outlined earlier, various techniques exist to investigate disease clustering. In order to meet the second research question, *"Is there evidence of disease clustering?"*, the location coefficient (LQ) for each TLA is calculated and visualised. Moreover, the nature of data available for this project allows the calculation of two autocorrelation models investigating whether potential clustering is statistically confirmed.

The location coefficient

The location coefficient describes the ratio between the observed and expected rate of a particular disease in a particular area over a specific period (Elliott and Wartenberg, 2004). This coefficient is frequently applied in order to investigate the occurrence of disease clustering, and is calculated using this formula (Walford, 1995):

$$LQ = \frac{X_i}{Y_i} : \frac{X}{Y}$$

X_i = observed rate of the specific disease in area i (TLAs),
Y_i = population in area i (TLAs),
X = observed rate of the specific disease in the overall area (New Zealand),
Y = population in the overall area (New Zealand).

Values greater than one represent local rates being above the average, whereas values of less than one indicate the rate being below the average. In this thesis this formula is applied as follows:

$$LQ = \frac{\text{rate per TLA}}{\text{average national rate}} .$$

Moran's I and Geary's c

The spatial autocorrelation models Moran's I and Geary's c are calculated to statistically confirm the occurrence of disease clustering within the data. Spatial autocorrelation is based on the principle that spatial data have a greater similarity to, or influence on, those locations (e.g. TLAs) within their immediate vicinity (National Institute of Justice, 2004; Wakefield et al., 2000). The global approach applied here is based on distance and adjacency methods respectively. Compared to the

Moran's I coefficient, the Geary's c is more sensitive for smaller spatial entities. Both models are calculated in order to control each of the statistical results. If the results of these models indicate the same tendencies, there is stronger evidence for the statistical validity of the findings.

The calculation of the Moran's I (I) and the Geary's c (C) is frequently applied in geographical epidemiology and based on these formulas (National Institute of Justice, 2004):

$$I = \frac{N \sum_i \sum_j W_{ij} (X_i - \bar{X})(X_j - \bar{X})}{(\sum_i \sum_j W_{ij}) \sum_i (X_i - \bar{X})^2}$$

$$C = \frac{(N-1) [\sum_i \sum_j W_{ij} (X_i - X_j)^2]}{2(\sum_i \sum_j W_{ij}) \sum_i (X_i - \bar{X})^2}$$

where N = number of cases, X_i = variable value at a particular location i, X_j = variable value at another location (where $I \neq j$), X = mean of the variable and W_{ij} = weight applied to the comparison between location i and location j.

The Moran's I varies between -1 and +1. If the variables are evenly distributed (negative values), this refers to negative spatial autocorrelation. Clustering is indicated by positive spatial autocorrelation (positive values), whereas values around zero are a sign of random distribution (no spatial autocorrelation). The Geary's c statistic reaches values between 0 and 2 where values about 1 indicate no spatial autocorrelation, values < 1 positive spatial autocorrelation, and values > 1 negative spatial autocorrelation. Both autocorrelation models are calculated with CrimStat®III (National Institute of Justice, 2004; Wakefield et al., 2000).

The standard deviation classification scheme

The last research question concerned with the analysis of the epidemiologic data investigates the appearance of outliers of campylobacteriosis across New Zealand: *Where are extreme values of campylobacteriosis located?* The standard deviation classification scheme (ESRI®ArcMap™ 9.1) is used to highlight areas deviating significantly from the mean in order to identify outliers. The computational method calculates the square root of the average squared deviation of all the data from the mean (Macfie and Nufrio, 2006). The two-colour ramp emphasises values above the mean (shown in blue) and below (shown in red).

4.2 Multiple regression in explanatory research

As the second aim, **"Gain insight into the relative importance of plausible determinants assumed to be affecting the distribution of campylobacteriosis in New Zealand"**, is of an explanatory nature, it is appropriate to apply regression analysis. Regression analysis is a major method used in geographic correlation studies in order to quantify statistical relationships. Assessing the effect of one independent on a dependent variable is called simple regression, whereas modelling numerous variables is called multiple regression (Macfie and Nufrio, 2006). Each regression model is based upon specific assumptions; for instance, the general linear modelling approach assumes the relationship between the dependent and independent variables being linear. Other assumptions are concerned with the normal distribution of the residuals, homoscedasticity[18], autocorrelation, and measurement errors (Macfie and Nufrio, 2006; Pedhazur and Schmelkin, 1991). In addition, it is possible to deduce spatial patterns from the analysis of the residuals and compare these results with the observed *Campylobacter* rates (Osborn, 2000).

Other regression techniques such as Poisson or logistic regression may be applied in geographic correlation studies. Hatch et al. (1991), for example, used logistic regression analysis to investigate the age-gender-specific probability of cancer in relation to proximity of residence to a plant. Another study (Williams et al., 2004) applied an ecological study design including GIS and Poisson regression to explore the geographic correlation between deprivation and risk of meningococcal disease.

Moreover, there are several ways of entering the independent variables in the modelling process. For example, the stepwise variable selection is based on the consideration of statistical significance. Only independent variables showing a statistically significant association with the dependent variable are included in the modelling process, and all other variables are automatically excluded from the model (Norušis, 2004). However, there are no scientific reasons for the levels set for statistical significance. Usually, a level of significance is reached at the 95% or 99% level. These boundaries are completely arbitrary and certainly do not include any context covering more than the statistical perspective (White, 2007). The hierarchical

[18] Homoscedasticity: the variance around the regression line is the same for all values of the independent variables (definition adopted from Macfie and Nufrio, 2006: 344).

method, further described in subsection 4.2.2, is based on theoretical reasons rather than solely on statistical significance and is popular in health related research (Blakely et al., 2006; Gottlieb et al., 2004; Leyland and Groenewegen, 2003; Weitzman et al., 2005).

For this analysis, it was preferred to apply a basic linear hierarchical modelling approach because *"unless a specific reason demands it, it would appear preferable to conduct most multivariate ecologic analyses as untransformed linear regressions"* (Rothman, 1986: 305). The regression models were calculated using SPSS 13.0 for Windows Integrated Student Version. Further, it is emphasised that the reasons for applying a statistical hypothesis test is *"the evaluation of the role of chance as one of many possible explanations for observed associations"* (Hill 1965, in Rothman, 1986: 20). The issues related to the interpretation of the results and the meaning of statistical significance in epidemiologic geographic research in particular are discussed later.

4.2.1 The theoretical model

Applying multiple regression in explanatory research requires the prior development of a theoretical model describing the interrelationships between the dependent and independent variables in order to establish background knowledge for the interpretation of the results (Greenland, 1998). Thus, deductive considerations rather than statistical measures are regarded as the crucial determinants for deciding about *"[…] the certainty with which a variable is considered a risk factor for diseases […]* (Miettinen and Cook, 1981 in Rothman, 1986: 92). Consequently, this approach can be used to research generally accepted hypotheses (Pedhazur, 1997; Rothman, 1986) such as associations concerning the epidemiology and ecology of campylobacteriosis.

The first research question of the second aim of the thesis is concerned with the development of the theoretical model: *Which plausible determinants that might advantage the appearance of campylobacteriosis are important in the local context?* The broad basis of existing studies allows some general assumptions about factors having an impact on the appearance of campylobacteriosis in New Zealand. In order to deduce this study's main hypotheses from the literature, it is reasonable to distinguish between these types of variables:

- variables describing environmental **conditions affecting the distribution of the organism** (climatic factors: temperature, rain, seasonality; landuse);

- variables describing **exposure** (contact with carriers, consumption of potentially contaminated food and water, occupational contact);
- variables describing **demographic characteristics** of the affected population: age, gender, ethnicity;
- variables describing **socioeconomic conditions** of the affected population (deprivation index);
- variables considering direct **notification artefacts**.

4.2.1.1 The independent variables in detail

Generally, the independent variables are grouped into three categories developed on the basis of the theoretical elaboration of section 3.3. The 23 independent variables chosen for this analysis are explained in detail following category 3 of the classification operationalising conditions that potentially contribute to the geographical variation in campylobacteriosis (table 5).

Table 5: Classification and description of the independent variables[19]

Variables					Description
Category 1	**Category 2**	**Category 3**	**Name**	**Unit per TLA**	
Dependent variable		Rate	R9705	per 100,000	crude Campylobacter rate 1997 - 2005
underlying epidemiology	conditions affecting the distribution of the organism	climate	Rain	mm	monthly mean maximum total rainfall
			Temp	℃	monthly mean maximum daily temperatures
		seasonal change	WIS	index	weighted index of interseasonal change
		landuse	Pastoral_1	%	landuse primarily pastoral (high and low producing grassland)
			Pastoral_2	%	landuse: scrub (e.g. fernland, gorse, broom, sub alpine shrubland); tussock grassland (e.g. low producing and depleted grassland)
			Pastoral_3	SU/sqkm	total stock units per sqkm, including dairy, beef, sheep, deer, pigs, goats, horses, poultry
	exposure	water	ECEX	%	intensity of drinking water quality monitoring: % of E. coli exceedances in drinking water
			RW	%	population with rainwater supply
			NAPWS	%	population not on registered water supply
			DWMM	per capita	number of drinking water measurements
		rurality	Rurality_1	per capita	total stock units per average population 97_01, including dairy, beef, sheep, deer, pigs, goats, horses, poultry
			Rurality_2	%	population living in highly rural/ remote area
		occupation	Occup	per 1,000	usual residence by occupation for the employed usually president population count, aged 15 years and over
		food	Rest	per 100,000	Number of restaurants, bars, tavernen, clubs with alcohol licence
			FaFo	per 100,000	number of fast food outlets
			FreFo	per 100,000	number of fresh food outlets
			SuMa	per 100,000	number of supermarkets
surveillance/ characteristics of affected population	demography	demographic characteristics of affected population particularly at risk	u15	%	people age < 15
			e25_44	%	people age 25 - 44
			male	%	males
			Europ	%	Europeans
	socioeconomic factors	deprivation	NZDep	index	New Zealand Deprivation Score 2001
	notification	artefacts	GP	per 100,000	number of general practitioners

(Data sources: appendix A2)

[19] For additional information see appendix A2 (description and sources of the original data) and A3 (data overview).

Climate

The climate data are derived from Landcare Research (Leathwick et al., 2002). These data were originally taken from publications of the former New Zealand Meteorological Service. Temperature (variable Temp) and rainfall (variable Rain) data cover the period from 1950 - 1980. The variables of average monthly rainfall and temperature were extracted from the Landcare database. Using spatial analyst tools, ESRI®ArcMap™ 9.1, the output file was converted from ASCII to Raster and aggregated to the TLA-level. The final variables represent the monthly mean maximum daily temperatures (°C) per TLA and monthly mean total maximum rainfall (mm) per TLA.

Seasonal Change

Within the scope of this thesis it is not possible to investigate the distribution of campylobacteriosis considering the different seasons separately. Therefore, an index representing the proportion of interseasonal change per TLA was developed from data published by Hearnden et al. (2003). Using the aggregated spring, summer, autumn, and winter rates for the years 1993 - 2000, the interannual change was calculated. In order to emphasise higher summer incidence and inter-seasonal variation in certain TLAs, the interannual change index was weighted, using the three types of seasonal grouping developed by Hearnden et al. (2003). The result is a weighted index of interseasonal change (WIS):

WIS = ∑(interannual change) x seasonal grouping.[20]

Landuse

The variables Pastoral_1 and Pastoral_2 were obtained from the New Zealand Landcover Database (LCDB2) which is a thematic classification of 42 land cover and landuse classes. The polygon features contain a code and boundary representing the land cover type for the period summer 1997 and summer 2001/ 02 (Thompson et al., 2003). Categories with intensive and extensive pastoral landuse were extracted, aggregated to the TLA-level (spatial analyst tools, ESRI®ArcMap™ 9.1), and converted into percentages.

[20] The exact calculation is presented in appendix A4. See also figure 9 and 10, chapter 3.

The variable Pastoral_3 represents the stock density (total stock units per square kilometre) for New Zealand at the TLA-level. The total number of livestock per TLA is jointly compiled by Statistics New Zealand (SNZ) and the New Zealand Ministry of Agriculture and Forestry (MAF, 2002a). In order to improve the comparability of different loads from livestock as a measurement of agricultural intensity, the total number of livestock was converted into stock units (SU). One SU equals 500 kg cattle, and the calculation uses these conversion units (BSES, 2006; MAF, 2006; The Lifestyleblock.co.nz, 2006):

- 1 dairy cattle = 1 SU,
- 1 horse = 0.8 SU,
- 1 beef cattle = 0.6 SU,
- 1 pig = 0.3 SU,
- each sheep, goat, deer = 0.2 SU,
- 1 poultry = 0.1 SU.

The total number of SU per TLA was then divided by the land area of New Zealand calculated in ESRI®ArcMap™ 9.1(spatial analyst tools):

Stock density = total number of SU per TLA / square km per TLA.

Water

The variables ECEX, DWMM, and NAPWS are environmental health indicators obtained from NZPHO (2006). Originally, NZPHO provides an environmental health indicator representing access to safe drinking water (% of people on registered water supply). In order to emphasise non access to safe drinking water per TLA, the percentage of people not on a registered water supply was calculated from the original data. The variable representing the population with rainwater supply per TLA in percent (RW) was calculated from data provided by WINZ and the Register of Community Drinking-Water Supplies in New Zealand (ESR, 2006a; ESR and MoH, 2006; MoH, 2006). The number of people living in communities with rainwater supply were extracted from this database, aggregated to the TLA-level and converted into percentages.

Rurality

Both rurality variables represent people's potential exposure to particular rural health hazards, such as increased contact with potentially infected animals or having a rainwater supply as the drinking water source. The variable Rurality_1 was calculated by dividing the total number of stock units by the population per TLA:

> *"Calculating stock units as a measure of agricultural intensity over a catchment in relation to [demographic] indicators offers a tool for assessing risk from potential waterborne zoonotic pathogens, [...]." (Till and McBride, 2005: 200).*

The variable Rurality_2 represents people living in highly rural/remote areas having a particularly high risk of becoming infected with campylobacteriosis. In the 2001 Census, highly rural areas were characterised as listed below (SNZ, 2006):

- Home to 76,449 people or 2% of New Zealand's usually resident population count. Population density of 0.5 people per km^2.
- Highest proportion of males to females (113.6 males per 100 females).
- Over half of the employed population (53.6%) worked in agriculture and fishery occupations, compared with 8.4% nationally.

The original data are provided by SNZ (2006) . According to various spatial, demographic and socioeconomic criteria, the urban/rural profile classifies areas in New Zealand into two main categories and seven subcategories:

- Urban areas: main urban areas, independent urban communities, satellite urban communities.
- Rural areas: rural areas with high urban influence, rural areas with moderate urban influence, rural areas with low urban influence, highly rural/ remote areas.

Each CAU in New Zealand is classified into one of these categories. The proportion of people living in each of these categories was aggregated to the TLA-level and converted into percentages. For example, in the Rodney District, 55.7% of the population lives in main urban areas, 10% in independent urban areas, 2.9% in satellite urban areas, 16.3% in rural areas with high urban influence, 7.1% in rural areas with moderate urban influence, 8% in rural areas with low urban influence, and 0% in highly rural/remote areas.

Occupation

The data representing people occupied in meat and dairy related professions (variable Occup) were provided and aggregated to the TLA-level by SNZ (SNZ, 2001).

Food

The food outlet data (variables Rest, FaFo, FreFo, SuMa) were provided by Pearce et al. (2006). The authors requested these data from the 74 TLAs, researched company websites and geocoded the precise location of each food outlet across New Zealand. The variable restaurant density (Rest) also incorporates bars, taverns and clubs with an alcohol licence because it is likely that these establishments also serve food. No dataset is older than 2002. For the scope of this thesis, these point data were aggregated to area data in ESRI®ArcMap™ 9.1 using the extension Hawth's Analysis Tools © 2002 - 2006, version 3.25.

Demographic characteristics of affected population particularly at risk

As mentioned earlier, it was not possible to use standardised rates of campylobacteriosis for this analysis. However, as demography is supposed to have a significant effect on the distribution of campylobacteriosis, this project considers the variables of age, gender, and ethnicity in its modelling approach. Demographic characteristics of the affected population particularly at risk of becoming infected are included as separate variables (u15, e25_44, male, Europ). All demographic data are available from SNZ (2002) and was converted into percentages per TLA.

The original and the combined age groups

There are two age groups (<15 and 25 to 44) supposed to have the highest risk of falling ill. This was explored in a correlation analysis describing the degree to which two variables are linearly related (Macfie and Nufrio, 2006). The original six age groups (< 5, 5 to 14, 15 to 24, 25 to 44, 45 - 64, >65) were correlated with the aggregated *Campylobacter* rate per TLA (1997 - 2005) using SPSS 13.0 for Windows Integrated Student Version. As expected, the three age groups <5, 5 to 15 and 25 to 44 were significantly associated with the *Campylobacter* rate. Table 6 presents the results of the correlation analysis.

Table 6: Degree of correlation between age and the *Campylobacter* rate per TLA, New Zealand 1997 - 2005

	original age groups						combined age group
	under 5	5_14	15_24	25_44	45_64	over 65	under 15
Pearson correlation coefficient	-0.272**	-0.488*	0.088	0.436*	0.093	-0.088	-0.437*

* Correlation is significant at the 0.01 level (2-tailed). ** Correlation is significant at the 0.05 level (2-tailed).

Interestingly, the age groups <5, and 5 to 15 were inversely correlated with the rate. The implication of this result is discussed later. For further analysis, these two age groups were combined into one age group, that is <5 to 15.

Deprivation

In New Zealand, socioeconomic disparities are measured as NZDep, the socioeconomic index of deprivation. The index consists of various individual variables describing socioeconomic disadvantage within the population (e.g. unemployment, low professional education, living conditions such as not living in own home). The index's ordinal scale ranges from 1 - 10, where 1 represents the areas with the least deprived scores and 10 the areas with the most deprived scores. This 10 point scale is derived from a first principal component score, which has been scaled to have a mean of 1,000 index points and a standard deviation of 100 index points. It has to be kept in mind that the deprivation score applies to areas rather than individual people.

Using the deprivation index at the TLA-level, which represents a relatively large spatial entity, provides the possibility of describing general tendencies, but it is not possible to draw conclusions for smaller spatial units such as census area units (CAU) or meshblocks (Salmond et al., 2006; Salmond and Salmond, 2002). NZDep01 is originally calculated for CAUs and meshblocks. Using the same methodology, the population weighted average deprivation score was calculated for the TLA-level (variable NZDep) (Salmond and Crampton, 2002: 11).

Notification

The general practitioner density (variable GP) was calculated using data from Pearce et al. (2006). The original dataset for their study was provided by the MoH. The data were collected in 2003 for the whole of New Zealand. 1,383 facilities, including the number of GPs, were recorded and geocoded. For this thesis, the ratio of general practitioners per population was calculated for the TLA level.

4.2.1.2 Hypotheses of the analysis

Due to the prior descriptions and explanations, the following is hypothesised:

Conditions affecting the distribution of the organism
Higher rates of campylobacteriosis appear in TLAs

- with a greater amount of rainfall (climate: Rain),
- with lower temperatures (climate: Temp),
- with a high weighted index of interseasonal change (seasonality: WIS).
- primarily characterised by rural landuse (landuse: Pastoral_1,Pastoral_2,Pastoral_3).

Exposure
Higher rates of campylobacteriosis appear in TLAs where the

- *E.coli* exceedances in drinking water are high (contaminated water: ECEX),
- proportion of people on rainwater supply is high (contaminated water: RW),
- proportion of people not on registered water supply is high (contaminated water: NAPWS),
- proportion of drinking water measurements is high (contaminated water: DWMM),
- stock density per population is high (rurality: Rurality_1),
- proportion of people living in rural areas is high (rurality: Rurality_2),
- proportion of people working in the meat and milk processing industry is high (occupational contact: Occup).
- restaurant density is high (contaminated food: Rest),
- fast food outlet density is high (contaminated food: FaFo),
- fresh food outlet density is high (contaminated food: FreFo),
- supermarket density is high (contaminated food: SuMa),

Demography
Higher rates of campylobacteriosis appear in TLAs with a high proportion of

- children under 15 and adults aged 24 to 45 (age: u15, e25_44),
- males (gender: male),
- Europeans (ethnicity: Europ).

Socioeconomic factors
Higher rates of campylobacteriosis appear in TLAs where the

- deprivation index is high (NZDep).

Notification
Higher rates of campylobacteriosis appear in TLAs where

- better access to health care (high GP-density) is provided (GP).

4.2.2 Hierarchical modelling

In hierarchical multiple regression the independent variables are entered in a predetermined order which is, for example, based on theoretical considerations (Cramer, 2003).[21] Usually, the modelling process *"start[s] with a model of minimal computational or conceptual complexity"* (Greenland, 1998: 403) and adds complexity with each level in order to assess each variable's contribution to the proportion of the variance in the dependent variable within different modelling steps (Cramer, 2003; Duncan et al., 1998).

For example, Blakely et al. (2006) investigated the association of neighbourhood-level volunteerism with mortality in New Zealand, focusing on how the association of neighbourhood volunteerism changes within different levels of the modelling process. They conducted a cohort study using multilevel Poisson regression analyses, including the following variables as plausible predictors for mortality in New Zealand:

- Model 1: neighbourhood volunteerism, baseline variables (ethnicity, marital status);
- Model 2: neighbourhood volunteerism, baseline variables, individual-level covariates (education, car access, labour force);
- Model 3: neighbourhood volunteerism, baseline variables, individual-level covariates, neighbourhood-level covariates (rurality, neighbourhood deprivation).

The first level, also called baseline regression model, adjusts only for demographic variables, whereas the second model includes the individual-level covariates. The final model incorporates all of the variables. Their results show that the association between neighbourhood volunteerism and mortality in New Zealand is not consistent within the modelling process. The final model shows no statistically significant independent association of a structural measure of neighbourhood volunteerism with mortality.

The thesis's modelling approach consists of five steps. Each level not only considers the relationship of the independent variables to the *Campylobacter* rate, but also takes into account how the variables are related to each other. The main interest is the analysis of how the effects of each of the variables and different sets of variables change with increasing complexity of the regression models. As hierarchical modelling offers the possibility of investigating the variables in their

[21] In the social science literature hierarchical modelling is also called multilevel modelling, and the different steps in the modelling process are called levels (Greenland, 1998).

complex interrelation, the approach is appropriate to assess the second and third research question of the thesis's second aim:

How are the identified determinants of campylobacteriosis related to the observed Campylobacter rate?

How much of the geographical variation in campylobacteriosis is explained by these factors?

Each modelling step is geared to a particular question emphasising the model's particular state of complexity based on theoretical considerations summarised in the variables' classification outlined earlier.

Step 1

How are the independent variables individually related to the *Campylobacter* rate?

The first modelling step starts with minimal complexity as it investigates each independent variable's individual relationship with the *Campylobacter* rate.

Step 2

How is each variable category three (e.g. climate/ interseasonal change) related to the *Campylobacter* rate, and how are the variables related to each other?

The second modelling step investigates how the observed associations of step one change when combining variables of the subcategories of climate/seasonal change, landuse, water, rurality/ occupation, food, and demographic characteristics of the affected population particularly at risk of becoming infected with campylobacteriosis.

Step 3

How is each variable category two (e.g. exposure) related to the *Campylobacter* rate, and how are the variables related to each other?

The third level increases the model's complexity, combining the categories of the former modelling step.

Step 4

How is each variable category one (e.g. underlying epidemiology) related to the *Campylobacter* rate, and how are the variables related to each other?

The fourth level further increases the model's complexity, combining the variables into the two main categories "underlying epidemiology" and "surveillance/ characteristics of the affected population".

Step 5

How are all variables related to the *Campylobacter* rate and to one another?

The final model is a summary of the data describing how all variables are linked; thus, it evaluates the independent variables' relative effects on the *Campylobacter* rate.

Reporting the results

The results report the R squares and the adjusted R squares, as well as the unstandardised *b* and the standardised *beta* coefficients. In contrast to the unstandardised *b* coefficients, the standardised *beta* coefficients are not affected by the scales being used (Pedhazur, 1997). The R square is the percent of the total sum of square that is explained. Since the R square can only increase, the adjusted R square can also decline in value if the contribution to the explained deviation by the additional variable is less than the impact on the degrees of freedom. Because the R square is a percent and the adjusted R square is referred to as an index value, it is preferred to mention the R square primarily. The coefficients *b* and *beta* and their respective p-values are reported in order to discuss the variables' indices of effects (*b*) on the dependent variable and the indices of the relative degree (*beta*) to which each of the variables contributes independently to the explanation of the dependent variable (Macfie and Nufrio, 2006; Pedhazur, 1997). The listing of the intercepts and the standard errors of the estimate complete the elements of the regression equation.

Analysing and mapping the residuals

The analysis of the residuals, or the unexplained variation of the dependent variable, gives information on the reliability of the regression model and is useful to identify areas of particular interest for further investigation. Generally, these are areas with high residual values where the model does not adequately explain variations in the dependent variable (Shaw and Wheeler, 1985). Mapping the relative or standardised size of residuals provides the possibility of visualising the magnitude of the variation between estimated and observed values. Standardised residuals are calculated by dividing the observed residual by the estimated standard error of the equivalent residual. Because the standard error of the estimate is a little larger than the standard deviation of the residuals, standardised residuals have a mean of 0 and a standard deviation of slightly less than one 1 (Norušis, 2004).

The examination of standardised residuals is a useful technique in geographical analysis and applied in various studies (Clark, 1967; Luke et al., 2000; Shaw and Wheeler, 1985). First, the residuals are tested in terms of the assumption of normality of the error term, spatial independency and homoscedasticity. Secondly, the standardised residuals of the final model (step 5) are mapped differentiating between positive (pink colour scheme) and negative deviations (blue colour scheme). Darker colours represent low residual values where the reliability of the regression model is relatively accurate. With fading colours, the explanatory character of the model is less precise.

4.3 Summary

First, the methods used to analyse the geographical distribution of the epidemiological data were described. Second, linear hierarchical modelling was introduced, based on the study's prior outlined theoretical assumptions. Moreover, the manipulation of the data and the underlying cartographical principles were discussed in detail, the latter also explaining reasons regarding the choice of class breaks and colour schemes. Theoretical considerations and the methodological approach determine the frame of reference in which the results of the study are discussed and evaluated.

5. Results

The first subsection presents the results referring to the study's first aim: **Examine the geographical distribution of campylobacteriosis in New Zealand.** The results are presented as choropleth maps, first showing the change of incidence and the average annual crude rate of campylobacteriosis in New Zealand for the period 1997 - 2005. Secondly, evidence of disease clustering is investigated by calculating the location coefficient and by applying two autocorrelation models. Finally, extreme values of campylobacteriosis are identified.

The second part illustrates the findings of investigating aim two: **Gain insight into the relative importance of plausible determinants assumed to be affecting the distribution of campylobacteriosis in New Zealand.** The results of each step of the regression analysis are presented in tables which report the associations of the explanatory variables and the *Campylobacter* rate as well as the statistical significance of the relationships. The residuals of the final regression model are analysed in detail and finally depicted as a choropleth map.

5.1 The geographical distribution of campylobacteriosis

A spatially differentiating view of the change of incidence in New Zealand for the period 1997 - 2005 reveals an increase of *Campylobacter* rates in 62 of 73 TLAs (85%) (figure 13). The highest increase appears in Nelson City (+377%), followed by the Wairoa (+363%) and Central Hawke's Bay District (+270%). 11 TLAs show a decrease of rates, headed by the Banks Peninsula District (-71%), followed by the Opotiki (-42%) and the Ruapehu District (- 36%). Spatial disparities in the overall trend of rising *Campylobacter* notifications in New Zealand are of particular interest in the course of this thesis as they might be explained by local particularities such as environmental, socioeconomic and demographic conditions.

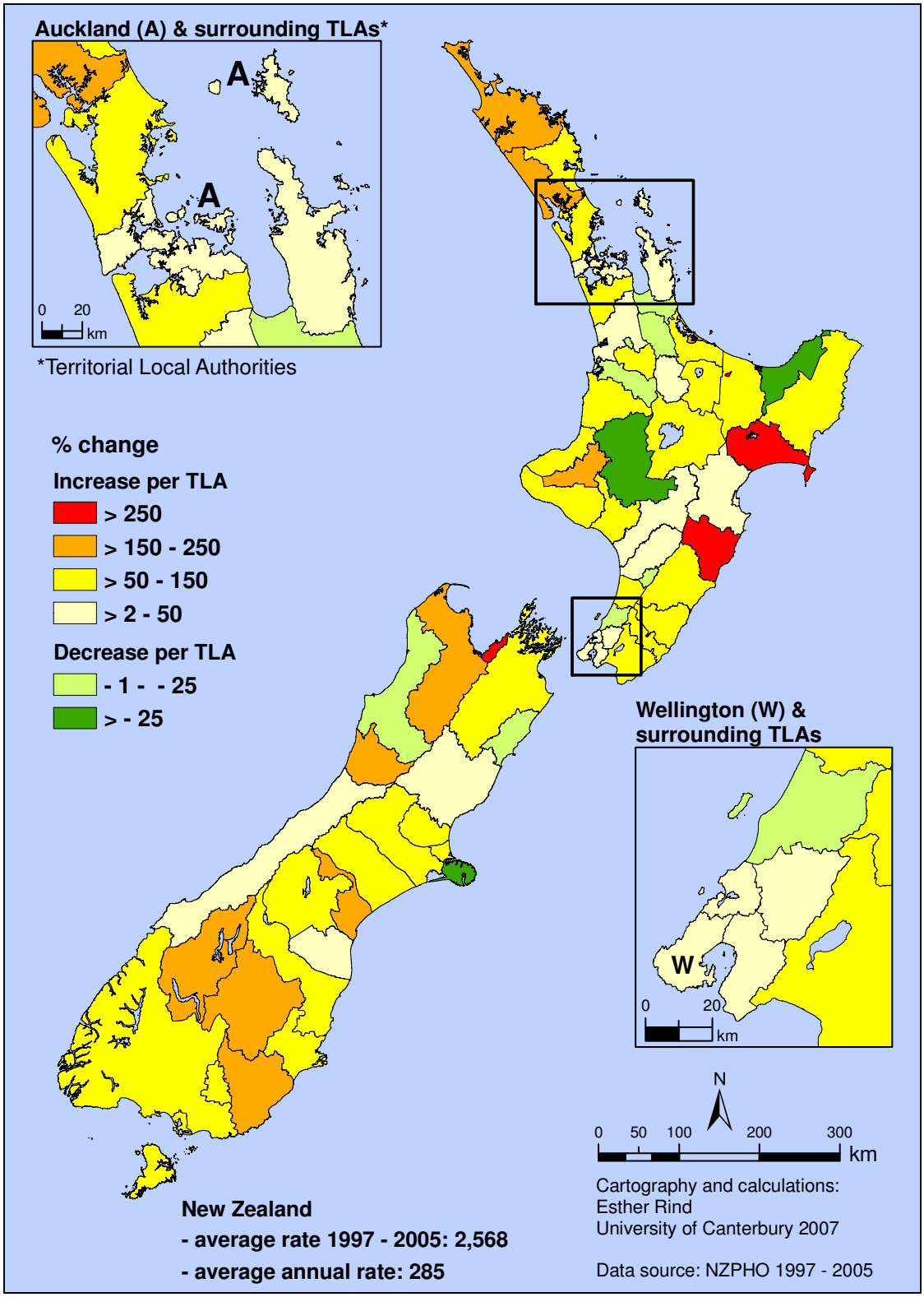


Figure 13: Campylobacteriosis in New Zealand – change between 1997 - 2005

Figure 14 displays the average annual crude rate of campylobacteriosis per TLA in New Zealand for the period 1997 - 2005.

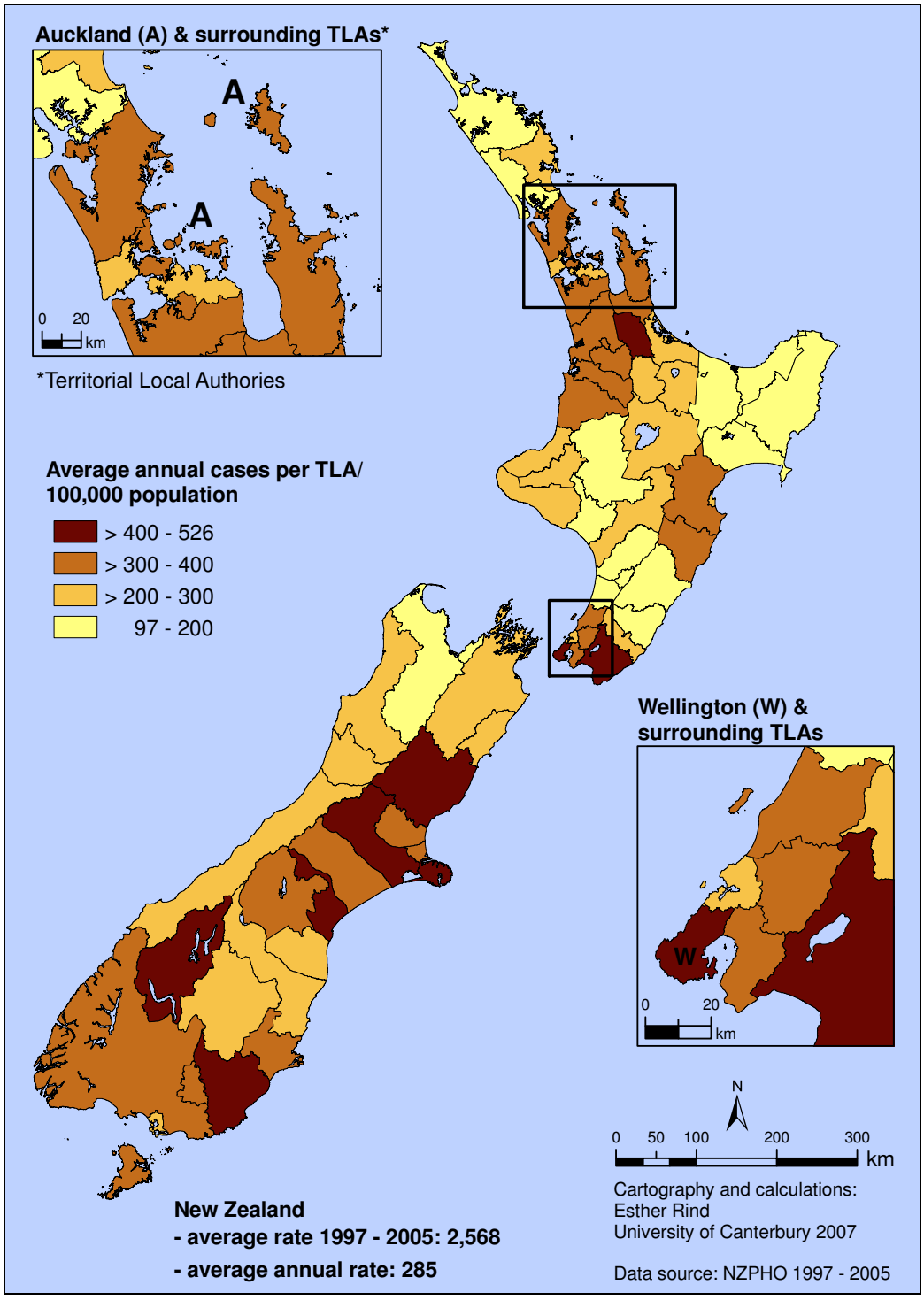


Figure 14: Campylobacteriosis in New Zealand 1997 - 2005 – average annual crude rate

New Zealand has rates of campylobacteriosis that are among the highest in the world, and the rates are not uniform across the country. In general, a large geographical variation in campylobacteriosis can be observed, ranging from an

average annual rate of 97/100,000 population (Opotiki District) to 526/100,000 population (Wellington City).[22]

Higher rates primarily occur in the South Island. There, 54% of the TLAs[23] are classified into the two upper classes (>300 - >400) whereas 41% of the North Island TLAs fall into the same category. The highest rate is observed for Wellington City (526). In the North Island, the highest rates (>400) are featured in TLAs around and south of Auckland as well as north and east of Wellington, primarily in the South Wairarapa District (451), the Matamata-Piako District (429) and Hamilton City (378). In the South Island, the highest rates (>400) are experienced in six TLAs: Selwyn (462), Queenstown-Lakes (458) Hurunui (3,945), Timaru (438), Clutha (419) and Banks Peninsula (412).

In the North Island, the lowest rate is exhibited in the Opotiki District (97), followed by the Far North (108) and the Kawerau District (115). In the South Island, rates in the lowest class (<200) are confined to two TLAs: Nelson City (149) and the Tasman District (168). Compared with the change in the rates for the period 1997 - 2005 (figure 13), these TLAs (except for the Opotiki district) interestingly show the highest increase. Further investigation of this observation reveals that lower rates per TLA are statistically significantly related to greater increase over time. The Pearson correlation coefficient between the rate (1997 - 2005) and the change of the rates over time is $r = -0.248$ ($p = 0.034$).

Additionally, these initial observations allow the preliminary conclusion that higher rates (>300 - >400) appear to be slightly clustered with a general north to south gradient and a concentration of higher rates in the South Island. This assumption is further investigated by the calculation of autocorrelation models and the location coefficient.

[22] All following rates are specified per 100,000 population.

[23] There are 49 TLAs in the North Island and 24 TLAs in the South Island (without Chatham Islands).

5.1.1 Disease clustering across New Zealand

This subsection deals with the investigation of disease clustering, *"a heterogeneous and clumped distribution of disease cases"* (Alexander and Cuzick 1992, in: Sabel and Löytönen, 2004: 52), in the *Campylobacter* rate in New Zealand for the period 1997 - 2005. First, the location coefficient for this period was calculated and visualised per TLA (figure 15).

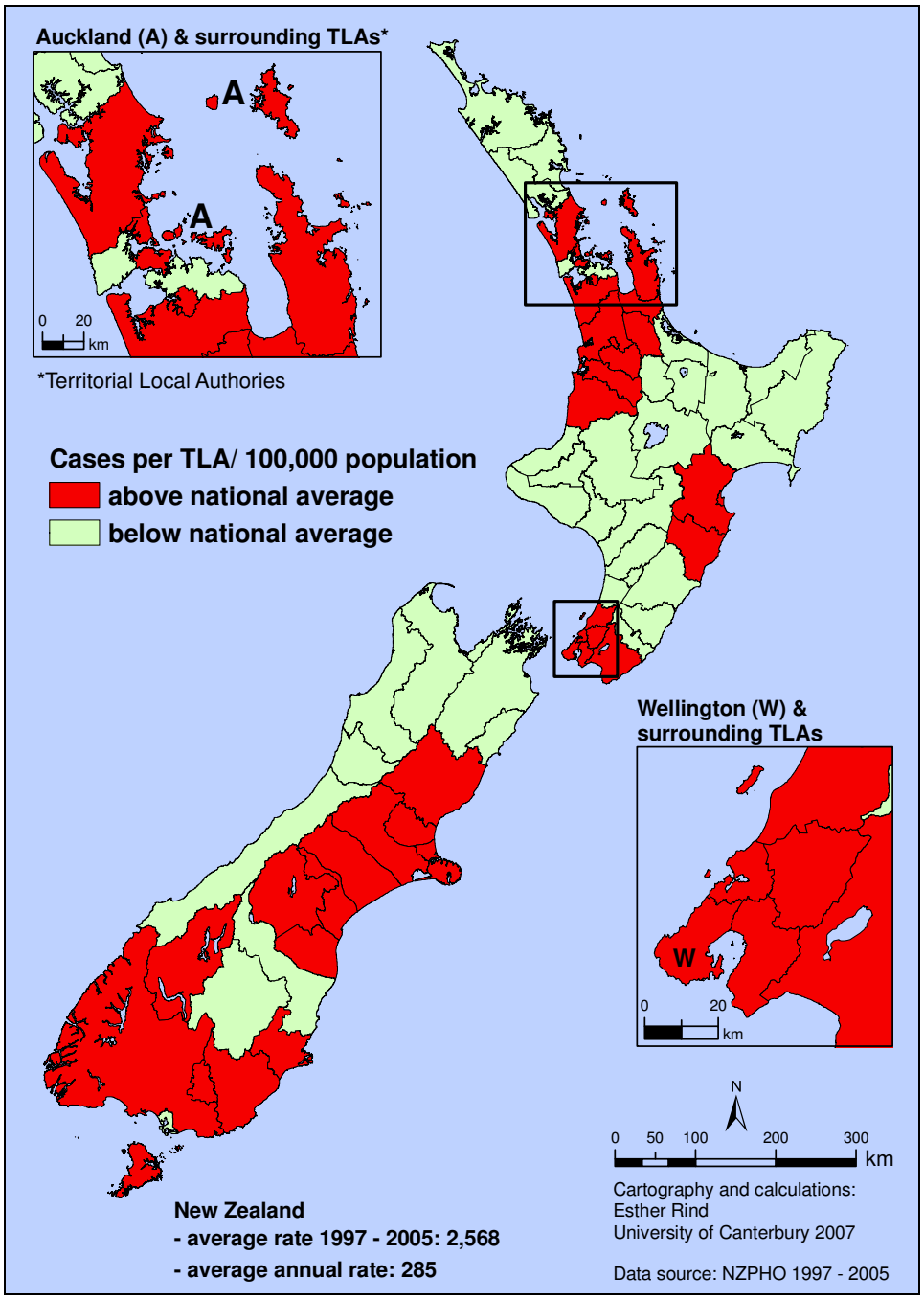


Figure 15: Campylobacteriosis in New Zealand 1997 - 2005 – location coefficient

This map compares patterns of disease cases above and below the national average and reveals striking regional clustering of rates above the annual national average

(285) across New Zealand. Distinct patterns of cases per TLA which are above the national average can be observed in three areas in the North Island and two areas in the South Island:

- North Island:
  - 13 TLAs north and south of Auckland and Auckland City,
  - 3 TLAs along the east coast,
  - 6 TLAs in the southern part of the island.
- South Island:
  - 9 TLAs along the east coast,
  - 5 TLAs in the southern part of the island.

This result is statistically confirmed with the calculation of the Geary's c and Moran's I. Both autocorrelation models show an apparent tendency of clustering (table 7).

Table 7: Autocorrelation: Geary's c and Moran's I

Autocorrelation model (Rate '97 - '05)	**Results** (all p-values are significant at 0.01)
Geary's c	0.85 (= positive spatial autocorrelation)
Moran's I	0.11 (= positive spatial autocorrelation)

The results further demonstrate that for both autocorrelation models there is less than 1% likelihood that this clustered pattern could be the result of random chance. However, as the Geary's c is close to 1 and the Moran's I close to 0 the statistical evidence of clustering is very low. This can be further investigated by visualising the *Campylobacter* rate using the standard deviation classification scheme presented in the next subsection.

5.1.2 The identification of outliers

The standard deviation (std) classification scheme is used to identify outliers of the *Campylobacter* rate of New Zealand for the period 1997 - 2005 per TLA (figure 16). Generally, there are 24 TLAs that present values distinctively deviating from the average rate (2,568). Of these, 15 are in the North Island and 9 are in the South Island. Statistical clustering for all outliers remains significant, but low (Moran's I = 0.2, $p < 0.01$).

In the North Island, Wellington is the only upper outlier (> +2 std). The Matamata-Piako District and the South Wairarapa District also show values positively varying from the mean (> +1 to +2 std). For the other 12 TLAs, however, the map illustrates

values varying negatively from the mean (< -1 to -2 std). These occur in the northern (Far North and Kaipara District), eastern (Whakateane, Kawerau, Opotiki and Wairoa District) and southern part (Manawatu, Tararua, Horowhenua, Masterton District and Palmerston North City) of the island as well as in the centre (Ruhapehu District).

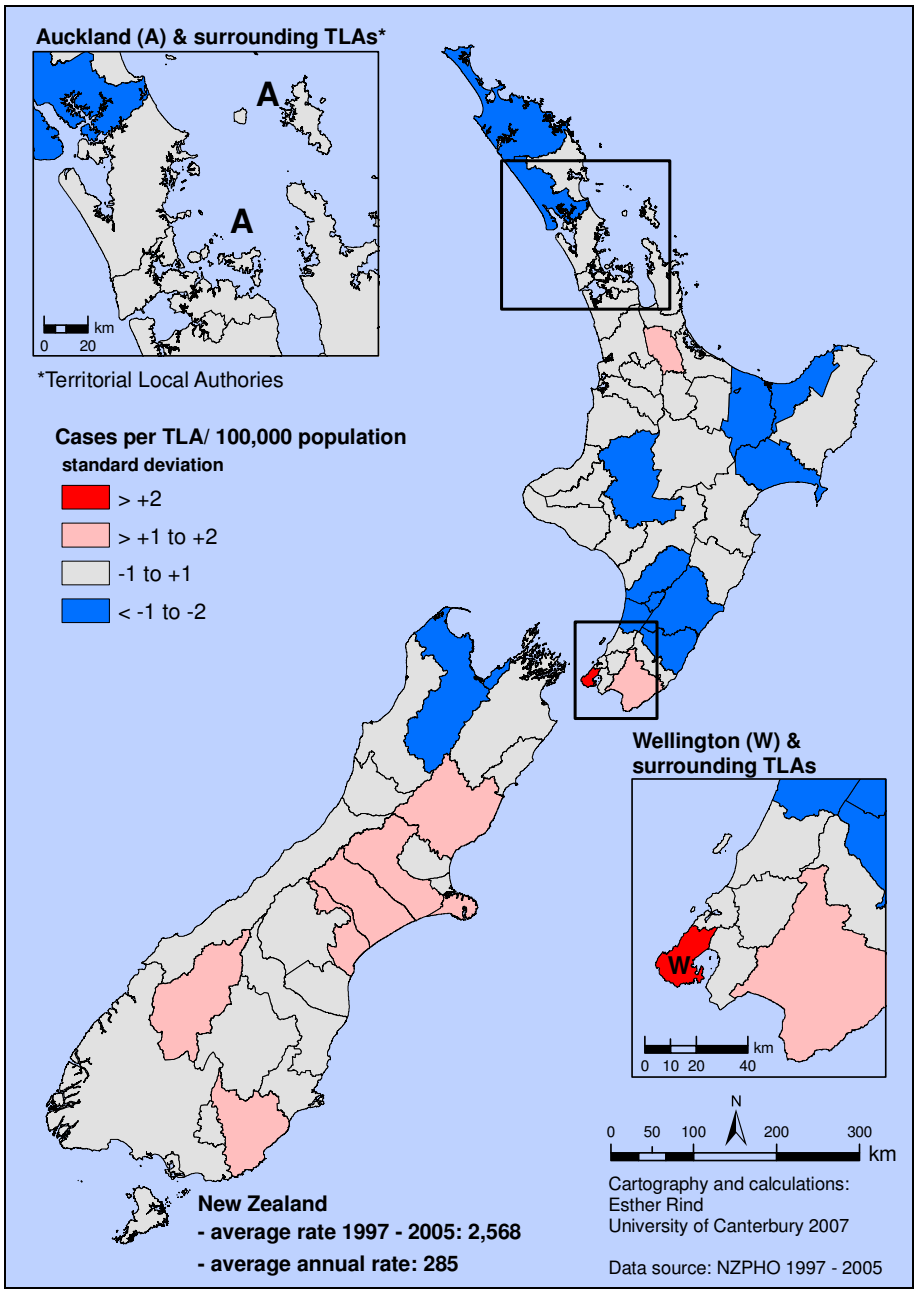


Figure 16: Campylobacteriosis in New Zealand 1997 - 2005 – extreme values

In the South Island, there are seven TLA showing values varying positively from the mean (> +1 to +2 std). Of these, five are concentrated along the east coast (Hurunui, Selwyn, Banks Peninsula, Ashburton and Timaru District). The other two TLAs, however, are not clustered (Queenstown-Lakes and Clutha District). Nelson City and the Tasman District are the only two TLAs in the South Island where the rate varies negatively from the mean (< -1 to -2 std).

5.2 Determinants of campylobacteriosis and the disease rate

This subsection presents the results of the statistical analysis which examine the relationship between the *Campylobacter* rate and plausible conditions contributing to the particular geographical occurrence of this disease. Table 8 summarises the minimum and maximum values, the mean and the standard deviation of all variables used for modelling.

Table 8: Descriptive statistics of the variables for TLAs

Variables					Descriptive Statistics			
Category 1	**Category 2**	**Category 3**	**Name**	**Unit per TLA**	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Dependent variable		*Campylobacter* rate	R9705	per 100,000	873.08	4735.11	2567.51	867.43
underlying epidemiology	conditions affecting the distribution of the organism	climate/ seasonal change	Rain	mm	115.00	1096.00	337.41	201.63
			Temp	℃	19.20	25.20	23.56	1.28
			WIS	index	3.54	285.37	55.87	55.60
		landuse	Pastoral_1	%	5.56	96.64	46.41	23.08
			Pastoral_2	%	<1.00	62.98	17.34	15.21
			Pastoral_3	SU/sqkm	1.59	340.05	87.83	66.13
	exposure	water	ECEX	%	<1.00	13.30	1.99	2.67
			RW	%	<1.00	12.54	1.31	2.31
			NAPWS	%	<1.00	52.80	16.78	14.63
			DWMM	per capita	<1.00	13.49	2.28	2.09
		rurality/ occupation	Rurality_1	per capita	<1.00	49.12	12.24	12.65
			Rurality_2	%	<1.00	72.76	5.66	13.33
			Occup	per 1,000	1.65	188.19	56.44	49.29
		food	Rest	per 100,000	<1.00	1044.20	156.14	163.30
			FaFo	per 100,000	5.71	174.03	73.70	28.55
			FreFo	per 100,000	38.77	318.28	114.14	48.45
			SuMa	per 100,000	<1.00	77.42	25.82	16.49
surveillance/ characteristics of affected population	demography	demographic characteristics of affected population particularly at risk	u15	%	18.12	31.28	23.50	2.76
			e25_44	%	23.74	36.73	28.23	2.42
			male	%	46.90	53.70	49.46	1.17
			Europ	%	42.17	93.05	75.65	13.12
	socioeconomic factors	deprivation	NZDep	index	913.57	1146.96	1000.13	45.55
	notification	artefacts	GP	per 100,000	14.58	67.23	36.57	10.89

Looking at the descriptive statistics it becomes clear that there is a large variation in the distribution of the *Campylobacter* rate and all independent variables. For example, the mean for the variable amount of rainfall per TLA (Rain) is 337.41mm ranging from 115mm in Invercargill City to 1,096mm in the Westland District. Referring to the stated hypothesis, a high amount of rainfall in a particular TLA could be associated with a higher *Campylobacter* rate, whereas a low amount of rainfall could be related to a lower disease rate. Accordingly, high and low C*ampylobacter* rates in different TLAs might be explained by higher and lower percentages of people particularly at risk of becoming infected (age groups under 15 and 25 to 44, males, Europeans). These disparities are analysed further in order to investigate whether there is statistical evidence explaining the geographical variation in campylobacteriosis in New Zealand.

5.2.1 The hierarchical models

Hierarchical models are used to assess the relationship between the independent variables and the *Campylobacter* rate (1997 - 2005). To begin with, each of the 23 explanatory variables are modelled separately. Then, the combined effect of variables classified in the categories three, two, and one are tested one after another. The final model assesses the interactions between the *Campylobacter* rate and all of the explanatory variables.

As outlined earlier, the assumptions necessary for the validity of the regression model are hardly fulfilled by geographic data. Here, the assumption of normality is violated for several of the independent variables. This is explored by executing the One-Sample Kolmogorov-Smirnov procedure, which tests the null hypothesis that a sample comes from a normal distribution. Large significance values (>.05) indicate that the observed distribution corresponds to the theoretical distribution (Norušis, 2004). Table 9 summarises the results of the test of normality for all variables.

Table 9: Results of the one-Sample Kolmogorov-Smirnov procedure

Variable	significance value	normal distribution
Rate_9705	0.941	yes
Rain	0.001	no
Temp	0.168	yes
WIS	0.001	no
Pastoral_1	0.606	yes
Pastoral_2	0.144	yes
Pastoral_3	0.231	yes
ECEX	0.001	no
RW	0.000	no
NAPWS	0.044	no
DWMM	0.022	no
Rurality_1	0.037	no
Rurality_2	0.000	no
Occup	0.019	no
Rest	0.008	no
FaFo	0.438	yes
FreFo	0.249	yes
SuMa	0.038	no
u15	0.984	yes
e25_44	0.507	yes
male	0.675	yes
Europ	0.155	yes
NZDep	0.603	yes
GP	0.555	yes

When the condition of normality is not true in the original data, this can often be achieved by applying a logarithmic transformation. However, the regression analysis was conducted using the original untransformed dataset as the *"need for data*

normality is subordinate" because *"transformation can lead to later interpretational difficulties"* (Shaw and Wheeler, 1985: 249).

5.2.1.1 Step 1: the univariate models

Generally, the univariate models explain up to 42% of the variance in the *Campylobacter* rate (table 10).

Table 10: Univariate associations of the explanatory variables and the *Campylobacter* rate per TLA, New Zealand 1997 - 2005

category 3	variable	R square	adj. R square	intercept	standard error of the estimate	*b* coefficient	p-value
climate/ seasonal change	Rain	0.009	-0.005	2,432.985	869.754	0.399	0.435
	Temp	0.060	0.047	6,497.193	846.776	-166.812	0.036**
	WIS	0.064	0.051	2,346.167	844.883	3.962	0.030**
landuse	Pastoral_1	0.001	-0.013	2,500.597	872.870	1.442	0.747
	Pastoral_2	0.063	0.050	2,318.944	845.481	14.336	0.032**
	Pastoral_3	0.016	0.002	2,444.894	853.343	1.638	0.289
water	ECEX	0.000	-0.014	2,573.371	873.447	-2.946	0.939
	RW	0.121	0.108	2,738.574	819.194	-130.454	0.003*
	NAPWS	0.007	-0.007	2,653.559	870.241	-5.129	0.467
	DWMM	0.013	-0.001	2,458.008	867.657	47.934	0.330
rurality/ occupation	Rurality_1	0.020	0.006	2,472.556	851.745	9.495	0.239
	Rurality_2	0.024	0.010	2,510.981	863.180	9.983	0.195
	Occup	0.016	0.002	2,443.066	866.929	2.205	0.291
food	Rest	0.015	0.001	2,668.065	867.069	-0.644	0.307
	FaFo	0.002	-0.012	2,459.617	872.498	1.464	0.686
	FreFo	0.026	0.013	2,236.040	861.947	2.904	0.170
	SuMa	0.018	0.004	2,386.219	865.692	7.022	0.260
demographic characteristics of affected population particularly at risk	u15	0.191	0.179	5,796.459	785.791	-137.416	0.000*
	e25_44	0.190	0.179	-1,844.832	786.032	156.298	0.000*
	male	0.006	-0.008	-361.374	870.747	59.211	0.504
	Europ	0.157	0.145	588.200	802.253	26.163	0.001*
deprivation	NZDep	0.417	0.409	14,865.987	666.966	-12.297	0.000*
artefacts	GP	0.033	0.020	3,097.569	858.924	-14.494	0.123

*statistical significance (sig) at the <0.01 level, **sig at the < 0.05 level

Separately modelled, the variables temperature (Temp), weighted index of interseasonal change (WIS) and one of the landuse variables (Pastoral_2) explain about 6% of the variance in the disease rate (adj. R^2 for these variables approximates 0.050). All of these regression coefficients are statistically significant ($p < 0.05$). The variable rainwater supply per population (RW) explains about 12% ($p < 0.05$) of the variation in campylobacteriosis (adj. $R^2 = 0.108$), whereas the variables referring to the demographic characteristics of the affected population (except for the variable of gender) explain between about 16% to 19% of the variation in the disease rate (adj. R^2 for both of the age variables = 0.179). All of these associations (except for the variable of gender) are statistically significant ($p < 0.01$). Entered separately, the variable NZDep explains the largest proportion

(42%) of the variation in campylobacteriosis (R^2 = 0.417, adj. R^2 = 0.409, p = 0.000). Each of the other variables explains less than 5% in the variation of the disease rate, and none of these associations are statistically significant.

Table 11 summarises the direction of the associations between the dependent and the independent variables when modelled separately and the directions expected according to the hypotheses which are based on the literature review. The table additionally reports the presence of statistical significance in these relationships.

Table 11: Hierarchical modelling step 1: associations and statistical significance

category 3	variable	direction of association		statistical significance
climate/ seasonal change	Rain	+	expected	no
	Temp	-	expected	**
	WIS	+	expected	**
landuse	Pastoral_1	+	expected	no
	Pastoral_2	+	expected	**
	Pastoral_3	+	expected	no
water	ECEX	-	inverse	no
	RW	-	inverse	*
	NAPWS	-	inverse	no
	DWMM	+	expected	no
rurality/ occupation	Rurality_1	+	expected	no
	Rurality_2	+	expected	no
	Occup	+	expected	no
food	Rest	-	inverse	no
	FaFo	+	expected	no
	FreFo	+	expected	no
	SuMa	+	expected	no
demographic characteristics of affected population particularly at risk	u15	-	inverse	*
	e25_44	+	expected	*
	male	+	expected	no
	Europ	+	expected	*
deprivation	NZDep	-	inverse	*
artefacts	GP	-	inverse	no

*sig at the <0.01 level, **sig at the < 0.05 level

For 16 of the 23 explanatory variables, the associations are in the anticipated direction. For example, it is hypothesised that the *Campylobacter* rate increases with a greater amount of rainfall and a higher WIS, but decreases with lower temperatures. These hypotheses are confirmed and statistically corroborated for both of the variables. However, for the other seven variables the direction of the association is the opposite to that expected, most notably, for all of the variables classified in the category water (except the variable DWMM), the age group <15, and

the deprivation index. Moreover, the associations of the variables RW, <15 and NZDep are statistically significant at the <0.05 level and the <0.01 level respectively.

5.2.1.2 Step 2: multivariate modelling of category 3 variables

Step 2 of the hierarchical modelling process assesses the multivariate associations of the combined explanatory variables classified in category 3. As both of the classes "deprivation" and "artefacts" incorporate only one variable already modelled in step one, step 2 of the modelling process consists of six separate models. The results are shown in table 12.

Table 12: Multivariate associations of the explanatory variables of category 3 and the *Campylobacter* rate per TLA, New Zealand 1997 - 2005

category 3	variable^	R square	adj. R square	intercept	standard error of the estimate	*b* coefficient	*beta* coefficient	p-value
climate/ seasonal change		0.097	0.058	5,424.991	842.020			
	Rain					-0.046	-0.011	0.931
	Temp					-128.171	-0.189	0.120
	WIS					3.173	0.203	0.113
landuse		0.136	0.097	1,874.314	811.496			
	Pastoral_1					-1.717	-0.047	0.795
	Pastoral_2					22.053	0.393	0.004*
	Pastoral_3					4.637	0.359	0.054***
water		0.143	0.092	2,541.494	826.467			
	ECEX					-4.409	-0.014	0.912
	RW					-157.125	-0.418	0.003*
	NAPWS					9.131	0.154	0.256
	DWMM					38.359	0.093	0.457
rurality/ occupation		0.027	-0.016	2,501.232	861.135			
	Rurality_1					9.970	0.148	0.623
	Rurality_2					5.568	0.087	0.573
	Occup					-1.163	-0.067	0.802
food		0.083	0.029	2,059.095	854.555			
	Rest					-1.414	-0.266	0.053***
	FaFo					2.374	0.078	0.557
	FreFo					3.302	0.184	0.206
	SuMa					6.867	0.131	0.336
demographic characteristics of affected population particularly at risk		0.389	0.353	-4,462.488	697.931			
	u15					-43.471	-0.138	0.421
	e25_44					157.862	0.441	0.000*
	male					38.952	0.052	0.647
	Europ					22.051	0.333	0.045**

*sig at the <0.01 level, **sig. at the < 0.05 level, *** close to sig at the < 0.05 level,
^ bold variables indicate multicollinearity **(mc)**

The first model climate/ seasonal change explains about 10% of the variation in the *Campylobacter* rate (adj. R^2 = 0.058). The partial regression coefficients (*b*) of the variables Rain and Temp indicate a negative association with the independent variable, whereas the relationship between the rate and the variable WIS is positive. This suggests that higher rates of campylobacteriosis are associated with a larger WIS. Lower disease rates, however, are probably linked to lower temperatures and

less rainfall per TLA. For example, with a one unit change in the variable Temp, the *Campylobacter* rate decreases by about 128 units assuming the other variables are held constant. Supposing no change in the values of the other variables, the variable WIS has the most significant individual effect on the disease rate (*beta* = 0.203) closely followed by the effect of the variable Temp (*beta* = -0.189). Both of the effects are approximately 20 times as great as the effect of the variable Rain (*beta* = -0.011). However, the p-values of the variables are not statistically significant.

The second model combines the three landuse variables and accounts for about 14% of the variation in the *Campylobacter* rate (adj. R^2 = 0.097). The partial correlation coefficient between the rate and the variable Pastoral_1 indicates a negative association, whereas the relationships for the variables Pastoral_2 and Pastoral_3 show a positive direction. The most significant *beta* coefficient is 0.393 (Pastoral_2) closely followed by the value of 0.359 (Pastoral_3). The association of the variable Pastoral_2 is statistically significant at the <0.01 level, whereas the variable Pastoral_3 is extremely close to reaching statistical significance at the <0.05 level (p = 0.054). This model also suggests issues concerned with multicollinearity. One way of detecting multicollinearity is to examine the correlation matrix of the independent variables (Appendix A5). Multicollinearity is indicated when high correlation coefficients occur between the independent variables. Tabachnick and Fidell (2001: 84 in Pallant, 2005) recommend not including two variables with a bivariate correlation close to r = 0.7 or higher in the same analysis. This model shows a high correlation coefficient (r = 0.771, p < 0.01) between the variables Pastoral_1 and Pastoral_3. This issue is further considered when modelling the final step of the hierarchical regressions (step 5).

The water related variables are combined in model four which explains about 14% in the variation of the dependent variable (adj. R^2 = 0.092). The regression coefficients of both of the variables ECEX and RW indicate a negative association with the disease rate, while the relationship for the other two variables NAPWS and DWMM is positive. From this model it can be concluded that TLAs with a low intensity of drinking water monitoring (ECEX) and less rainwater supply per population (RW) experience lower rates of campylobacteriosis; conversely, the higher the proportion of people not on registered water supply (NAPWS) and the higher the number of drinking water measurements per capita (DWMM) the more

likely a higher disease rate per TLA. The variable RW shows the highest and only statistically significant standardised regression coefficient (*beta* = -0.418, $p < 0.01$).

Model four incorporates the three variables representing the category rurality and occupation respectively. This model explains about 3% percent of the variation in the *Campylobacter* notifications (adj. R^2 = -0.016). The positive direction of both of the variables Rurality_1 and Ruality_2 indicate higher disease rates with increasing stock units per population (Rurality_1) and a higher percentage of people living in remote areas (Rurality_2). The model also suggests lower rates of campylobacteriosis in TLAs with a high number of people employed in dairy and meat related industries (Occup). The variable Rurality_1 indicates the highest independent effect on the disease rate (*beta* = 0.148), but none of the explanatory variables shows a statistically significant association with the independent variable. Moreover, the relationship between the variables Rurality_1 and Occup indicates multicollinearity ($r = 0.886$, $p < 0.01$).

The four food related variables are combined in the fifth model which accounts for about 8.3% of the variation in the *Campylobacter* rate (adj. R^2 = 0.029). All of the variables, except the variable representing restaurant density (Rest), have a positive association with the dependent variable. This suggests that a higher fast food, fresh food and supermarket density per TLA is associated with higher rates of campylobacteriosis. A higher restaurant density, however, is related to a lower disease rate. The variable Rest has the highest and only significant *beta* coefficient of all variables (*beta* = -0.266). None of the variables is significantly related to the dependent variable, but the p-value of the variable Rest is extremely close to reaching statistical significance at the 0.05 level ($p = 0.053$).

The final model explains approximately 39% of the variation in the *Campylobacter* rate (adj. R^2 = 0.353). All variables, except the age group under 15 (u15), are positively related to the disease rate. This indicates that TLAs with a greater percentage of people aged 25 to 44, males and Europeans face higher disease rates. Both of the associations of the variables age 25 to 44 (e25_44) and Europeans (Europ) are statistically significant at the <0.01 level and the <0.05 level respectively. The variable e25_44 shows the highest independent effect on the dependent variable (*beta* = 0.441). Moreover, the variables u15 and Europ are significantly related with each other ($r = -.712$, $p < 0.01$).

Table 13 summarises and evaluates these results with respect to the study's hypotheses.

Table 13: Hierarchical modelling step 2: associations and statistical significance

category 3	variable^	direction of association		statistical significance
climate/ seasonal change	Rain	-	invers	no
	Temp	-	expected	no
	WIS	+	expected	no
landuse	**Pastoral_1**	-	invers	no
	Pastoral_2	+	expected	*
	Pastoral_3	+	expected	***
water	ECEX	-	invers	no
	RW	-	invers	*
	NAPWS	+	expected	no
	DWMM	+	expected	no
rurality/ occupation	**Rurality_1**	+	expected	no
	Rurality_2	+	expected	no
	Occup	-	invers	no
food	Rest	-	invers	***
	FaFo	+	expected	no
	FreFo	+	expected	no
	SuMa	+	expected	no
demographic characteristics of affected population particularly at risk	**u15**	-	invers	no
	e25_44	+	expected	*
	male	+	expected	no
	Europ	+	expected	**

*sig at the <0.01 level, **sig at the < 0.05 level,
***close to sig at the < 0.05 level **^mc**

For 14 of the 21 explanatory variables, the associations are in the anticipated direction. As mentioned earlier, it is hypothesised that the *Campylobacter* rate increases with a greater amount of rainfall and a higher WIS, but decreases with lower temperatures. When these variables are modelled together, the hypotheses are confirmed for both of the variables Temp and WIS. The variable Rain, however, changes its prior positive univariate association with the dependent variable to a negative relationship. Likewise, a change in the direction of associations is observed for the variables Pastoral_1, NAPWS and Occup. The level of statistical significance also changes for the variables Temp, WIS, Pastoral_2, Pastoral_3, Rest, u15, and Europ. For example, both of the variables Temp and WIS show statistically significant associations with the *Campylobacter* rate when modelled separately. However, when these variables are incorporated in the climate/ seasonal change model, the associations become statistically insignificant.

5.2.1.3 Step 3: multivariate modelling of category 2 variables

Step 3 of the modelling approach assesses the multivariate associations of the combined explanatory variables classified in category 2. As there is no further differentiation for the already modelled categories “demography”, “socioeconomic factors” and “notification”, step 3 of the multivariate modelling consists of two models. The results are shown in table 14.

Table 14: Multivariate associations of the explanatory variables of category 2 and the *Campylobacter* rate per TLA, New Zealand 1997 - 2005

category 2	variable^	R square	adj. R square	intercept	standard error of the estimate	*b* coefficient	*beta* coefficient	p-value
		0.166	0.089	4,239.914	815.104			
conditions affecting the distribution of the organism	Rain					-0.085	-0.020	0.879
	Temp					-97.195	-0.145	0.244
	WIS					1.595	0.104	0.436
	Pastoral_1					-2.307	-0.063	0.743
	Pastoral_2					16.560	0.295	0.053***
	Pastoral_3					4.464	0.346	0.072***
		0.254	0.117	2,407.604	802.556			
exposure	ECEX					-55.054	-0.173	0.241
	RW					-154.150	-0.419	0.004*
	NAPWS					0.882	0.015	0.927
	DWMM					-19.359	-0.048	0.728
	Rurality_1					-0.439	-0.007	0.983
	Rurality_2					12.709	0.199	0.245
	Occup					2.234	0.129	0.621
	Rest					-1.043	-0.200	0.142
	FaFo					0.038	0.001	0.993
	FreFo					4.209	0.240	0.113
	SuMa					0.297	0.006	0.973

*sig at the <0.01 level, *** close to sig at the < 0.05 level, **^mc**

The first model incorporates variables of the category conditions affecting the distribution of the organism and explains about 17% of the variation in the *Campylobacter* rate (adj. R^2 = 0.089). The unstandardised regression coefficients of the variables Rain, Temp and Pastoral_2 show a negative relationship to the disease rate. Thus, it can be suggested that less rainfall, lower temperatures and less pastoral landuse inversely affect the geographical distribution of campylobacteriosis. The other three variables (WIS, Pastoral_2, Pastoral_3) have a positive association with the disease rate, indicating, for example, that TLAs with a high WIS experience higher rates of campylobacteriosis. The standardised regression coefficient of the variable Pastoral_3 (*beta* = 0.346) has the most significant individual effect on the dependent variable. None of the associations are statistically significant, but the p-values of the variables Pastoral_2 (p = 0.053) and Pastoral_3 (p = 0.072) are very close to reaching statistical significance at the <0.05 level. Further, the association

between the variables Pastoral_1 and Pastoral_3 indicate multicollinearity ($r = 0.771$, $p < 0.01$).

The second model includes variables representing the category exposure and accounts for about 25% of the variation in the disease rate (adj. $R^2 = 0.117$). The *b* coefficients of the variables ECEX, RW, DWMM, Rurality_1 and Rest indicate a negative association with the distribution of campylobacteriosis, whereas the other six variables (NAPWS, Rurality_2, Occup, FaFo, FreFo, SuMa) are positively related to the disease rate. The variable RW shows the highest and only statistically significant *beta* coefficient (*beta* = -0.419, $p < 0.01$) suggesting that the variable representing the number of people with rainwater supply has the largest independent effect on the disease rate. Multicollinearity is indicated for the association between the variables Rurality_1 and Occup ($r = .714$, $p < 0.01$).

These results are summarised and evaluated in table 15.

Table 15: Hierarchical modelling step 3: associations and statistical significance

category 2	variable^	direction of association		statistical significance
conditions affecting the distribution of the organism	Rain	-	inverse	no
	Temp	-	expected	no
	WIS	+	expected	no
	Pastoral_1	-	inverse	no
	Pastoral_2	+	expected	***
	Pastoral_3	+	expected	***
exposure	ECEX	-	inverse	no
	RW	-	inverse	*
	NAPWS	+	expected	no
	DWMM	-	inverse	no
	Rurality_1	-	inverse	no
	Rurality_2	+	expected	no
	Occup	+	expected	no
	Rest	-	inverse	no
	FaFo	+	expected	no
	FreFo	+	expected	no
	SuMa	+	expected	no

*sig at the <0.01 level, ***close to sig at the < 0.05 level, **^mc**

For 10 of the 17 variables the association is in the expected direction. Comparing the associations between this and the prior modelling approach there is a change from a positive to a negative association for the variables DWMM and Rurality_1. The opposite effect can be observed for the variable Occup. Moreover, there is a change in the level of significance for the variables Pastoral_2 and Rest as both of the variables' associations become distinctively statistically insignificant.

5.2.1.4 Step 4: multivariate modelling of category 1 variables

The last but one step of the hierarchical modelling process combines explanatory variables classified in category 1. The results of both models are shown in table 16.

Table 16: Multivariate associations of the explanatory variables of both categories 1 and the *Campylobacter* rate per TLA, New Zealand 1997 – 2005

category 1	variable^	R square	adj. R square	intercept	standard error of the estimate	*b* coefficient	*beta* coefficient	p-value
underlying epidemiology		0.375	0.178	3,560.808	774.595			
	Rain					0.132	0.031	0.837
	Temp					-72.169	-0.107	0.433
	WIS					1.919	0.126	0.398
	Pastoral_1					3.654	0.099	0.663
	Pastoral_2					17.936	0.320	0.114
	Pastoral_3					3.568	0.276	0.162
	ECEX					-92.174	-0.290	0.079***
	RW					-124.116	-0.337	0.024**
	NAPWS					0.427	0.007	0.965
	DWMM					8.202	0.020	0.883
	Rurality_1					-5.259	-0.078	0.809
	Rurality_2					20.910	0.328	0.072***
	Occup					-0.090	-0.005	0.986
	Rest					-0.683	-0.131	0.350
	FaFo					-2.558	-0.084	0.541
	FreFo					4.442	0.253	0.110
	SuMa					-6.333	-0.123	0.487
surveillance/ characteristics of affected population		0.481	0.434	14,060.621	652.467			
	u15					-31.955	-0.102	0.563
	e25_44					71.233	0.199	0.114
	male					-1.867	-0.003	0.982
	Europ					-9.889	-0.150	0.480
	NZDep					-11.711	-0.615	0.005*
	GP					-5.477	-0.069	0.499

*sig at the <0.01 level, **sig. at the < 0.05 level, *** close to sig at the < 0.05 level, **^mc**

The model underlying epidemiology explains 37.5% of the variance in the *Campylobacter* rate (adj. R^2 = 0.178). Nine of the 17 partial regression coefficients are positively associated with the disease rate. For example, with a one unit change in the variable ECEX, the *Campylobacter* rate decreases by approximately 82 units. The other eight variables show a positive relationship to the disease rate. For instance, from this model it can be concluded that TLAs with a high percentage of people living in remote areas (Rurality_2) experience a higher rate of campylobacteriosis. The variables Pastoral_2, Rurality_2 and RW have the greatest independent effects on the disease rate (*beta* > 0.300). The latter is also statistically significant at the < 0.05 level, whereas the p-values of both of the variables ECEX and Rurality_2 are very close to being significant at the < 0.05 level. Multicollinearity is indicated for the variables Pastoral_1 and Pastoral_3 (r = .771, p < 0.01), and for the variables Rurality_1 and Occup (r = 0.886, p < 0.01).

The second model including variables representing factors of surveillance and characteristics of the affected population accounts for about 48.1% of the variation in the *Campylobacter* rate (adj. R^2 = 0.434). All of the model's six variables indicate a negative association with the disease rate, except for the variable representing the age group 25 to 44 (e25_44). So, it can be suggested that TLAs with a high percentage of people aged 25 to 44 face higher rates of campylobacteriosis. The variable NZDep has the greatest independent effect on the disease rate assuming the other variables are held constant (*beta* = -0.615). This association is also statistically significant at the <0.01 level. Further, the variable u15 is highly correlated with both of the variables Europ (r = -.712, p < 0.01) and NZDep (r = .678, p < 0.01), and the variable Europ is also significantly associated with the deprivation index (r = -.791, p < 0.01).

Table 17 summarises the results of the hierarchical modelling step 4.

Table 17: Hierarchical modelling step 4: associations and statistical significance

category 1	variable^	direction of association		statistical significance
underlying epidemiology	Rain	+	expected	no
	Temp	-	expected	no
	WIS	+	expected	no
	Pastoral_1	+	expected	no
	Pastoral_2	+	expected	no
	Pastoral_3	+	expected	no
	ECEX	-	invers	***
	RW	-	invers	**
	NAPWS	+	expected	no
	DWMM	+	expected	no
	Rurality_1	-	invers	no
	Rurality_2	+	expected	***
	Occup	-	invers	no
	Rest	-	invers	no
	FaFo	-	invers	no
	FreFo	+	expected	no
	SuMa	-	invers	no
surveillance/ characteristics of affected population	**u15**	-	invers	no
	e25_44	+	expected	no
	male	-	invers	no
	Europ	-	invers	no
	NZDep	-	invers	*
	GP	-	invers	no

*sig at the <0.01 level, **sig at the < 0.05 level,
***close to sig at the < 0.05 level **^mc**

For 11 of the 23 independent variables, the associations are in the anticipated direction. The relationships of the dependent and independent variables of RW and NZDep are statistically significant at the <0.01 level and the < 0.05 level respectively, whereas the associations of the variables ECEX and Rurality_2 are very close to

reaching statistical significance at the < 0.05 level. Compared to the prior modelling steps, there is a change in the direction of the associations from negative to positive for the variables Rain, Pastoral_1 and DWMM. The opposite effect can be observed for the variables Occup, FaFo, SuMa, male and Europ. The prior associations of the variables Pastoral_2 and Pastoral_3 with the disease rate, which were close to statistical significance at the < 0.05 level, become statistically insignificant. The same effect can be observed for the variables e25_44 and Europ. Further, the relationship of the variable ECEX becomes almost significant at the <0.05 level. The same effect can be observed for the variable Ruralitiy_2. Moreover, the level of significance of the variable RW changes from being significant at the <0.01 level to the <0.05 level.

5.2.1.5 Step 5: the final multivariate model

The final Model A incorporates all 23 independent variables, but with respect to the prior mentioned issue of multicollinearity Model B excludes independent variables having a bivariate correlation close to r = 0.7.

Model A

Table 18 presents the multivariate associations of all explanatory variables of the combined category 1 and the *Campylobacter* rate.

Table 18: Multivariate associations of all explanatory variables of the combined category 1 and the *Campylobacter* rate per TLA, New Zealand 1997 – 2005

	variable^	R square	adj. R square	intercept	standard error of the estimate	*b* coefficient	*beta* coefficient	p-value
		0.615	0.429	18,742.428	645.495			
	Rain					0.073	0.017	0.894
	Temp					-1.299	-0.002	0.988
	WIS					2.602	0.170	0.188
	Pastoral_1					3.695	0.100	0.606
	Pastoral_2					-6.967	-0.124	0.515
	Pastoral_3					1.224	0.095	0.593
	ECEX					-3.429	-0.011	0.943
	RW					-79.601	-0.216	0.099***
underlying epidemiology/ surveillance/ characteristics of affected population	NAPWS					-2.687	-0.049	0.748
	DWMM					1.984	0.005	0.966
	Rurality_1					-8.046	-0.119	0.672
	Rurality_2					10.644	0.167	0.291
	Occup					3.362	0.194	0.483
	Rest					-0.442	-0.085	0.472
	FaFo					-5.948	-0.196	0.126
	FreFo					4.498	0.256	0.059***
	SuMa					6.916	0.134	0.412
	u15					-100.246	-0.307	0.176
	e25_44					113.620	0.324	0.055***
	male					-142.191	-0.195	0.328
	Europ					-24.455	-0.360	0.147
	NZDep					-8.151	-0.409	0.093***
	GP					-12.726	-0.163	0.166

*** close to sig at the < 0.05 level, **^mc**

The final model A explains about 61.5% of the variability in the *Campylobacter* rate (adj. R^2 = 0.429) and of the 23 explanatory variables, 10 are positively associated with the disease rate. Each of the variables RW, FreFo, u15, e25_44, Europ and NZDep show independent effects on the dependent variable greater than *beta* = 0.200 and *beta* = -0.200 respectively. None of the model's variables indicate statistical significance, but the p-values of the variables RW, FreFo, e25_44 and NZDep are very close to reaching statistical significance at the <0.05 level. The issues concerned with multicollinearity are complemented by the significant correlation between both of the variables occup and male (r = .714, p < 0.01).

Table 19 summarises and evaluates the results with respect to the variables' directions of association and their statistical significance.

Table 19: Hierarchical modelling step 5A: associations and statistical significance final model A

	variable^	direction of association		statistical significance
underlying epidemiology/ surveillance/ characteristics of affected population	Rain	+	expected	no
	Temp	-	expected	no
	WIS	+	expected	no
	Pastoral_1	+	expected	no
	Pastoral_2	-	inverse	no
	Pastoral_3	+	expected	no
	ECEX	-	inverse	no
	RW	-	inverse	***
	NAPWS	-	inverse	no
	DWMM	+	expected	no
	Rurality_1	-	inverse	no
	Rurality_2	+	expected	no
	Occup	+	expected	no
	Rest	-	inverse	no
	FaFo	-	inverse	no
	FreFo	+	expected	***
	SuMa	+	expected	no
	u15	-	inverse	no
	e25_44	+	expected	***
	male	-	inverse	no
	Europ	-	inverse	no
	NZDep	-	inverse	***
	GP	-	inverse	no

***close to sig at the < 0.05 level **^mc**

For 11 of the 23 variables, the relationships are in the expected direction. Comparing this modelling step to the prior one, there is a change in the direction of associations and their statistical significance for some of the variables. The direction of association changes for the variables Pastoral_2 and NAPWS from positive to negative. The reverse effect can be observed for the variables Occup and SuMa.

Further, the prior almost statistically significant association of the variables ECEX and Rurality_2 become distinctively insignificant; the opposite effect is found for the variables FreFo and e25_44. Both of the variables RW and NZDep, however, alter their prior significant associations at the <0.05 level and <0.01 level to those almost reaching statistical significance.

Model B

Table 20 presents the multivariate associations of explanatory variables of the combined category 1 and the *Campylobacter* rate excluding variables showing high bivariate correlation coefficients; those are: Pastoral_1, Pastoral_3, Rurality_1, Occup, u15, male, Europ and NZDep. It was preferred to exclude the variable Pastoral_1 rather than the variable Pastoral_3 as the direction of the association of the latter remains stable during the modelling process. Because the variable Rurality_1 is highly correlated with both of the variables Occup and male, and the variable Occup is also highly correlated with the variable male, the variables Rurality_1 and Occup were dropped from this model. Moreover, both of the variables' associations are very unstable during the modelling process. The variables u15 and Europ were also excluded from Model B since both of them are highly correlated with the variable NZDep. Further, the variable u15 is significantly correlated with the variable Europ. The implications of these decisions are discussed and evaluated later.

Table 20: Multivariate associations of particular explanatory variables of the combined category 1 and the *Campylobacter* rate per TLA, New Zealand 1997 – 2005

	variable^	R square	adj. R square	intercept	standard error of the estimate	*b* coefficient	*beta* coefficient	p-value
underlying epidemiology/ surveillance/ characteristics of affected population		0.583	0.441	12,978.522	638.669			
	Rain					-0.205	-0.049	0.664
	Temp					12.991	0.019	0.868
	WIS					2.721	0.178	0.144
	Pastoral_2					-4.439	-0.079	0.651
	Pastoral_3					1.680	0.130	0.296
	ECEX					-6.921	-0.022	0.883
	RW					-65.361	-0.177	0.143
	NAPWS					-4.283	-0.073	0.616
	DWMM					0.794	0.002	0.986
	Rurality_2					12.206	0.191	0.193
	Rest					-0.436	-0.084	0.455
	FaFo					-4.211	-0.139	0.247
	FreFo					4.107	0.234	0.076***
	SuMa					4.174	0.081	0.600
	e25_44					129.410	0.369	0.005*
	male					-143.890	-0.197	0.184
	NZDep					-7.105	-0.357	0.007*
	GP					-10.693	-0.137	0.212

*sig at the <0.01 level, *** close to sig at the < 0.05 level

Hence, the final model B incorporates 18 instead of 23 explanatory variables and explains about 58% of the variation in the *Campylobacter* rate (adj. R^2 = 0.441). Of the 18 included explanatory variables, eight are positively related to the disease rate. The variable e25_44 has the highest independent effect on the distribution of campylobacteriosis (*beta* = 0.369), closely followed by both of the variables NZDep (*beta* = -0.357) and FreFo (*beta* = 0.234). The variables e25_44 and NZDep are significant at the <0.01 level, whereas the association of the variable FreFo is very close to reaching statistical significance at the <0.05 level. None of these 18 independent variables is highly correlated with one another; thus, this model corrects for the issue of multicollinearity.

Table 21 summarises the results of the final hierarchical modelling step.

Table 21: Hierarchical modelling step 5B: associations and statistical significance final model B

	variable^	direction of association		statistical significance
underlying epidemiology/ surveillance/ characteristics of affected population	Rain	-	inverse	no
	Temp	+	inverse	no
	WIS	+	expected	no
	Pastoral_2	-	inverse	no
	Pastoral_3	+	expected	no
	ECEX	-	inverse	no
	RW	-	inverse	no
	NAPWS	-	inverse	no
	DWMM	+	expected	no
	Rurality_2	+	expected	no
	Rest	-	inverse	no
	FaFo	-	inverse	no
	FreFo	+	expected	***
	SuMa	+	expected	no
	e25_44	+	expected	*
	male	-	inverse	no
	NZDep	-	inverse	*
	GP	-	inverse	no

*sig at the <0.01 level, ***close to sig at the < 0.05 level

For seven of 18 variables the associations are in the anticipated direction. Comparing these results to those of Model A, the directions of both of the associations of the climate variables Rain and Temp are in the opposite direction; all other relationships remain constant. In concordance with Model A, the level of significance for the variable FreFo remains stable, almost reaching statistical significance at the <0.05 level. The association of the variable RW, however, becomes statistically insignificant, whereas both of the variable's relationships e25_44 and NZDep reach statistical significance at the <0.01 level.

5.2.2 The residuals of the final multivariate model

The residuals of the regression analysis can be investigated and visualised in order to explain patterns of interrelated phenomena such as the geographical distribution of campylobacteriosis and its plausible determinants. First, the residuals are checked in terms of the assumption of normality of the error term. The shape of the histogram follows approximately the shape of the normal curve and is acceptable close to the normal curve (figure 17A). Further, the P-P plotted residuals follow relatively accurately a 45-degree line (figure 17B). The scatterplot (figure 17C) shows good scatter as the points fall randomly in a band around 0 to ± 2 (Norušis, 2004).

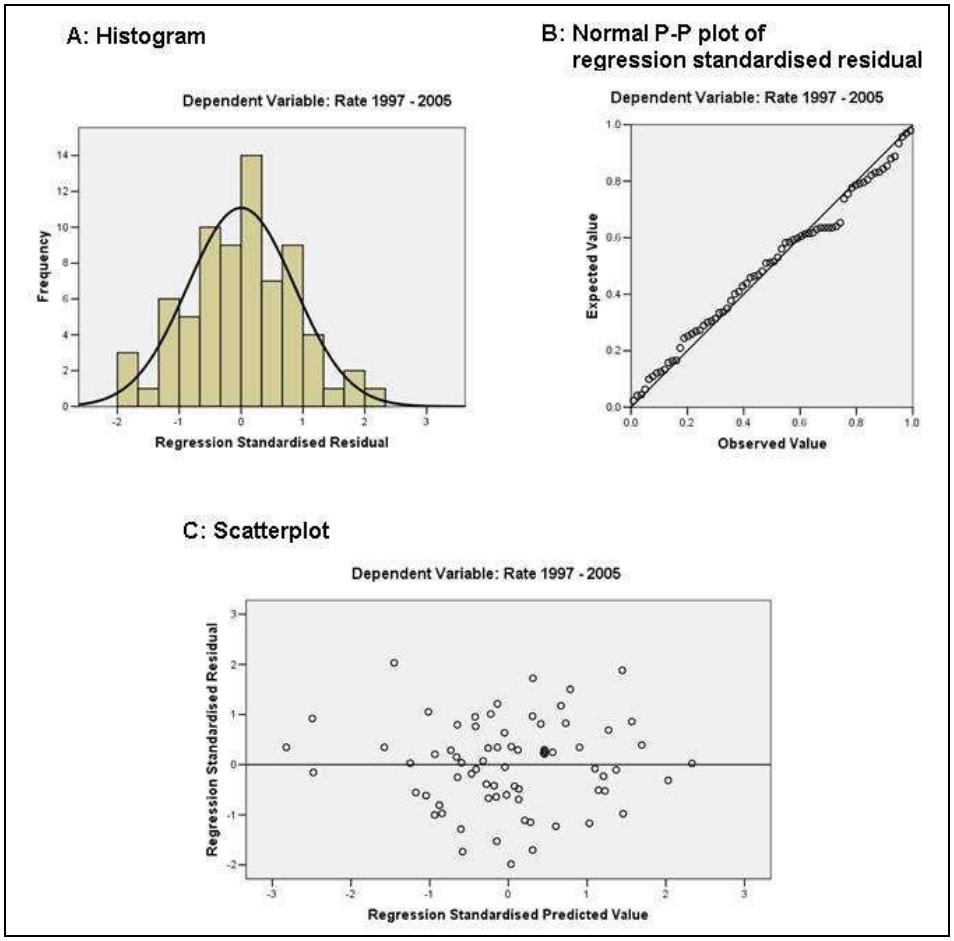


Figure 17: Histogram, P-P plot and scatter plot of the standardised residuals

Second, the One-Sample Kolmogorov-Smirnov procedure indicates the distribution of the residuals resembling a normal distribution (sig = 0.808). Finally, the Durbin-Watson statistic, a test for autocorrelated residuals, was performed as one of the assumptions of regression analysis is that the residuals for observations are uncorrelated. If this is true, the expected value of the Durbin-Watson statistic is 2. Values less than 2 indicate positive autocorrelation; values greater than 2 indicate negative autocorrelation. The results of the test are very close to two

(Durbin-Watson = 2.254). Consequently, none of these tests indicates that the assumption of normality of the error term is seriously violated (Norušis, 2004).

Figure 18 presents the standardised residuals of the final Model B from the multivariate hierarchical linear regression of the determinants of campylobacteriosis in New Zealand 1997 - 2000. It is apparent that the residuals are both positive and negative. Moreover, the negative residuals are more spatially concentrated: the Moran's I for the positive residuals is -0.059 (random distribution), whereas the Moran's I for the negative residuals is 0.183 ($p < 0.01$), meaning that there is less than 1% likelihood that this clustered pattern could be the result of random chance.[24]

The first part of the map visualises 38 TLAs showing positive standardised residuals. Of these, 30 TLAs are part of the first two residual classes (< 0.50 to 1.00) where the model explains the appearance of campylobacteriosis relatively well. Seven TLAs are confined to the residual classes three to four (1.01 to > 2.00), where the model explains the occurrence of the disease with decreasing accuracy. The best statistical explanation of the variation in campylobacteriosis is found for the Queenstown-Lakes District; there, the deviance between the observed and the expected value is very close to 0 (residual = 0.025). The most significant positive deviation, which represents also the only upper outlier (residual > 2.00), occurs for the Thames-Coromandel District (residual = 2.032).

The second part of the map displays the negative standardised residuals. Overall, there are 34 TLAs having negative residuals, of which 24 explain the occurrence of campylobacteriosis relatively well (residual classes -1.00 to < -0.50). The other two classes (residual classes -1.50 to -2.00) show 10 TLAs where the appearance of campylobacteriosis is explained with decreasing accuracy. There are no lower outliers (residuals > -2.00). The most significant negative deviation appears in Nelson City (residual = -1.985).

[24] The test of autocorrelation can be directly executed in ESRI® ArcMap™ 9.1 once the data are integrated into an ArcMap project. Therefore, the Moran's I of the residuals is calculated using the spatial statistics toolbox.

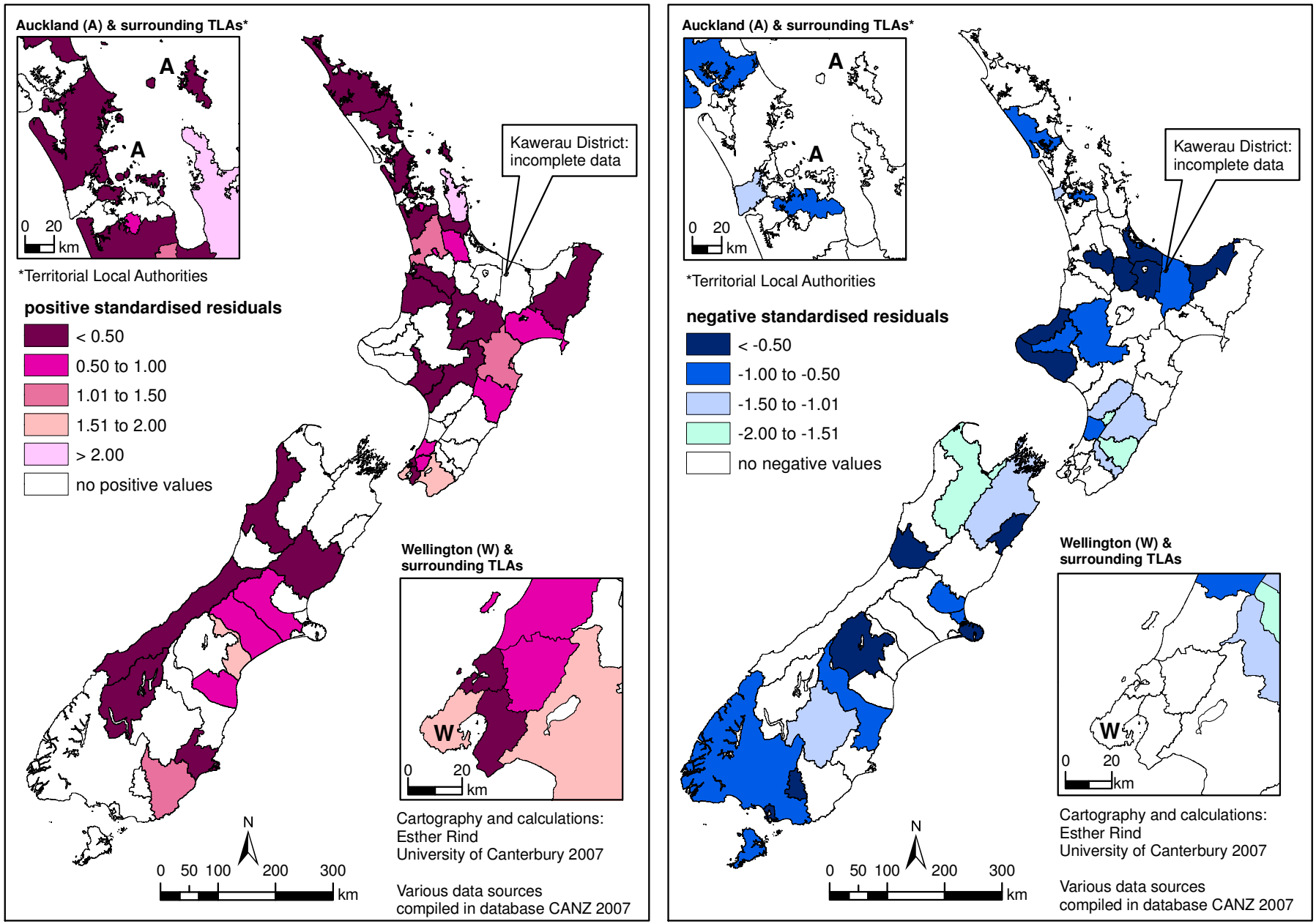


Figure 18: Standardised residuals from the hierarchical linear regression of the determinants of campylobacteriosis, New Zealand 1997- 2005

The residuals for the Kawerau District cannot be displayed as there are no data available to calculate the variables Pastoral_3 and Rurality_1. Generally, the Kawerau District experiences a rather low *Campylobacter* rate (1,039) compared to the average rate (2,568). Lower rates are present only in the Opotiki District (873) and the Far North District (969). Comparing the respective value of each of the independent variables of the Kawerau District to those of the other districts and cities (see Appendix A3), most of the values of the Kawerau District are confined to the lower third of values, meaning that the theoretical risk of becoming ill is relatively low compared to the values of the other TLAs. For example, compared to the other TLAs, the Kawerau District has the lowest percentage of Europeans and the second lowest fast food density. This listing summarises the relative position of each of the 23 independent variables of the Kawerau District:

- not available: Pastoral_3, Rurality_1;
- lower third (= low theoretical risk of becoming infected): Rain, Pastoral_2, RW, NAPWS, Occup, Rest, FaFo, FreFo, e25_44, Europ;
- medium third (= medium theoretical risk of becoming infected): WIS, Pastoral_1, Rurality_2, SuMa;
- upper third (= high theoretical risk of becoming infected): Temp, ECEX, DWMM, u15, male, NZDep, GP.

5.3 Summary

Key findings of the analysis are summarised according to the aims and research questions of the thesis.

Examine the geographical distribution of campylobacteriosis in New Zealand.

How is campylobacteriosis distributed across New Zealand?

There is a large geographical variation in campylobacteriosis across New Zealand, ranging from an average annual rate of 97/100,000 (Opotiki District) to 526/100,000 (Wellington City). Lower rates per TLA for 1997 - 2005 are statistically significantly related to greater increase over time.

Is there evidence of disease clustering?

Generally, there is statistical evidence for global and local clustering of the disease rate. Local clustering is observed for three areas in the North Island and two areas in the South Island.

Where are extreme values of campylobacteriosis located?

Upper and lower outliers of campylobacteriosis occur in New Zealand; however, higher rates primarily appear in the South Island.

Gain insight into the relative importance of plausible determinants assumed to be affecting the distribution of campylobacteriosis in New Zealand.

Which plausible determinants that might advantage the appearance of campylobacteriosis are important in the local context?

The theoretical model based on the literature incorporates 23 independent variables (or plausible determinants of campylobacteriosis) that might affect the distribution of campylobacteriosis in New Zealand. These variables are classified into five main categories; those are conditions affecting the distribution of the organism, exposure, demography, socioeconomic factors, and notification.

How are the identified determinants related to the observed Campylobacter rate?

The regression models show three kinds of associations between the disease rate and the explanatory variables throughout the modelling process:

- The relationships are stable and show the expected associations.
- The relationships are stable and show inverse associations.
- The relationships are unstable.

How much of the geographical variation of campylobacteriosis is explained by these factors?

Generally, the univariate models explain between < 1% and > 41% of the variance in the *Campylobacter* rate. For 16 of the 23 explanatory variables, the associations are in the anticipated direction. The highest R^2 is calculated for the variable NZDep ($R^2 = 0.417$; adj. $R^2 = 0.409$).

Modelled separately, the demographic variables (deprivation, age, ethnicity) explain the most significant proportion of the variation in campylobacteriosis (16% - 42%). About 6% - 12% of the variation in campylobacteriosis is explained by variables of the categories water (RW), climate (WIS, Temp) and landuse (Pastoral_2). The remaining variables explain less that 6% of the variation in campylobacteriosis.

Modelling the multivariate associations of the explanatory variables of category 3 and the *Campylobacter* rate, the model “demographic characteristics of the affected population particularly at risk” explains the most significant proportion of the variation in campylobacteriosis (38.9%), followed by the models of “water” (14.3%), “landuse” (13.6%), climate/seasonal change (9.7%), “food” (8.3%), and “rurality/occupation” (2.7%). The variable representing the age group 25 to 44 shows the most statistically significant independent effect on the variation in campylobacteriosis (*beta* = 0.441) and the variable amount of rainfall (Rain) the least independent effect (*beta* = -0.011).

Modelling the multivariate associations of the explanatory variables of category 2 and the *Campylobacter* rate, the model “exposure” explains 25.4% of the variation in campylobacteriosis, whereas the model “conditions affecting the distribution of the organism” explains 16.6% of the variation. The variable representing the percentage of people supplied by rainwater (RW) shows the most statistically significant

independent effect on the variation in campylobacteriosis (*beta* = -0.419) and the variable fast food density (FaFo) the least independent effect (*beta* = 0.001).

Modelling the multivariate associations of the explanatory variables of category 1 and the *Campylobacter* rate, the model "underlying epidemiology" explains 37.5% of the variation in campylobacteriosis, and the model "surveillance/characteristics of affected population" explains 48.1% of the variation in the disease rate. The variable deprivation (NZDep) shows the most statistically significant independent effect on the variation in campylobacteriosis (*beta* = -0.615) and the variable representing the percentage of males per TLA (male) the least independent effect (*beta* = -0.003).

The final model A of the hierarchical regression explains 61.5% of the variation in campylobacteriosis in New Zealand (adj. R^2 = 0.429). The final model B corrects for the issue of multicollinearity and explains 58.3% of the variation in the rate (adj. R^2 = 0.441). The variable representing the percentage of people aged 25 to 44 (e25_44) shows the most statistically significant independent effect on the variation in campylobacteriosis (*beta* = 0.369), closely followed by the statistically significant association of the variable NZDep (*beta* = -0.357). The variable representing fresh food outlet density (*beta* = 0.234), however, is very close to reaching statistical significance. Further, the variable representing the number of drinking water measurements per capita (DWMM) shows the least independent effect on the variation in the disease rate (*beta* = 0.002).

The analysis of the residuals of the hierarchical modelling shows both positive and negative residuals. The model explains the variation in campylobacteriosis relatively well for 75% of all TLAs (54 TLAs from 72 TLAs, without the Chatham Islands and the Kawerau district).

Table 22 summarises key results of the univariate and multivariate modelling which are discussed and evaluated in the next chapter.

Table 22: Summary of the univariate and the multivariate modelling

		step 1: univariate models			step 2: multivariate models category 3			step 3: multivariate models category 2			step 4: multivariate models category 1			step 5: final model A			step 5: final model B		
category 3	**variable**	**direction of association**		**sig**	**direction of association**		**sig**	**category 2**	**direction of association**		**sig**	**category 1**	**direction of association**		**sig**	**direction of association**		**sig**	**direction of association**		**sig**
climate/ seasonal change	Rain	+	expected	no	-	invers	no	**conditions affecting the distribution of the organism**	-	inverse	no	**underlying epidemiology**	+	expected	no	+	expected	no	-	inverse	no
	Temp	-	expected	**	-	expected	no		-	expected	no		-	expected	no	-	expected	no	+	inverse	no
	WIS	+	expected	**	+	expected	no		+	expected	no		+	expected	no	+	expected	no	+	expected	no
landuse	Pastoral_1	+	expected	no	-	invers	no		-	inverse	no		+	expected	no	+	expected	no			
	Pastoral_2	+	expected	**	+	expected	*		+	expected	***		+	expected	no	-	inverse	no	-	inverse	no
	Pastoral_3	+	expected	no	+	expected	***		+	expected	***		+	expected	no	+	expected	no	+	expected	no
water	ECEX	-	inverse	no	-	invers	no	**exposure**	-	inverse	no		-	invers	***	-	inverse	no	-	inverse	no
	RW	-	inverse	*	-	invers	*		-	inverse	*		-	invers	**	-	inverse	***	-	inverse	no
	NAPWS	-	inverse	no	+	expected	no		+	expected	no		+	expected	no	-	inverse	no	-	inverse	no
	DWMM	+	expected	no	+	expected	no		-	inverse	no		+	expected	no	+	expected	no	+	expected	no
rurality/ occupation	Rurality_1	+	expected	no	+	expected	no		-	inverse	no		-	invers	no	-	inverse	no			
	Rurality_2	+	expected	no	+	expected	no		+	expected	no		+	expected	***	+	expected	no	+	expected	no
	Occup	+	expected	no	-	invers	no		+	expected	no		-	invers	no	+	expected	no			
food	Rest	-	inverse	no	-	invers	***		-	inverse	no		-	invers	no	-	inverse	no	-	inverse	no
	FaFo	+	expected	no	+	expected	no		+	expected	no		-	invers	no	-	inverse	no	-	inverse	no
	FreFo	+	expected	no	+	expected	no		+	expected	no		+	expected	no	+	expected	***	+	expected	***
	SuMa	+	expected	no	+	expected	no		+	expected	no		-	invers	no	+	expected	no	+	expected	no
demographic characteristics of affected population particularly at risk	u15	-	inverse	*	-	invers	no					**surveillance/ characteristics of affected population**	-	invers	no	-	inverse	no			
	e25_44	+	expected	*	+	expected	*						+	expected	no	+	expected	***	+	expected	*
	male	+	expected	no	+	expected	no						-	invers	no	-	inverse	no	-	inverse	no
	Europ	+	expected	*	+	expected	**						-	invers	no	-	inverse	no			
deprivation	NZDep	-	inverse	*									-	invers	*	-	inverse	***	-	inverse	*
artefacts	GP	-	inverse	no									-	invers	no	-	inverse	no	-	inverse	no

Background:

dark grey: the relationships are stable and show the expected associations throughout the modelling process

light grey: the relationships are stable and show inverse associations throughout the modelling process

white: the relationships are unstable throughout the modelling process

6. Discussion

The results of the analysis are discussed, compared and evaluated in relation to the aims of the thesis and the context of national and international findings. The first aim of this thesis was to examine the distribution of campylobacteriosis across New Zealand. The analysis of the notification data (1997 - 2005) established the spatial and primarily descriptive background for this study. The second aim investigated plausible explanations for the observed spatial patterns of the first part of the analysis. The six research questions provided the structural basis for different steps of the analysis, whereas the discussion of the results links the separate findings in order to provide a comprehensive presentation of plausible explanations for the distribution of campylobacteriosis across New Zealand.

6.1 The distribution of campylobacteriosis across New Zealand

The *Campylobacter* rates of 1997 - 2005 show an increase over time for nearly all TLAs and a large geographical variation in campylobacteriosis. These patterns are not only observed in New Zealand, but consistent with those found in several other affluent nations. For example, research from the UK (Louis et al., 2005), Norway (Sandberg et al., 2006), Sweden (Nygård et al., 2004) and Australia (Morgan, 2002) describe a continuous increase of *Campylobacter* rates and differences in the distribution of campylobacteriosis since the 1990. Furthermore, the results suggest that higher rates of campylobacteriosis appear in the South Island rather than in the North Island. This pattern is consistent with many previous studies (e.g. ESR, 2006b; Mitchell et al., 2002).

Temporal patterns of campylobacteriosis are usually investigated focusing on the seasonality of the disease (Eyles et al., 2003; Hearnden et al., 2003; Kovats et al., 2005; Nylen et al., 2002) or the spatiotemporal development of disease clustering (Morgan, 2002). The result of this analysis adds to these findings in terms of providing statistical evidence for a relationship between time and the extent of rates. The observed association is indeed small, though statistically significant. The result follows a pattern earlier described by Kleinschmidt et al. (2002): the lower the rates in the first place, the more likely a higher increase over time. The findings of their study show the highest increase in incidence occurring where rates had previously been the lowest. They conclude that there has been some degree of stabilisation of rates in areas with the lowest initial incidence. In their study, this is explained by

environmental (e.g. climatic particularities) and population-related factors (e.g. immunity, demography) limiting the increase in rates in some areas. These explanations are also relevant in the context of this analysis.

6.2 Plausible determinants of campylobacteriosis

The investigated plausible determinants of campylobacteriosis that might be particularly relevant in the New Zealand context were examined in tandem, and the key findings from the hierarchical regression analysis, including the analysis of the residuals, are discussed here.

6.2.1 Conditions affecting the distribution of the organism

Climate and seasonal change

The variables representing climatic conditions (rainfall, temperature) show unstable associations with the *Campylobacter* rate throughout the modelling process. As soon as further variables (e.g. landuse, water) were added to the subsequent models, the relationships show changing directions of association, indicating that they are not independently associated with the *Campylobacter* rate. For example, the parameter estimate for the association between rainfall and the disease rate almost changes in each modelling step; thus, this association is highly related to all other identified determinants of campylobacteriosis. The result is in line with previous research. Kovats et al. (2005), for example, could not find a strong effect of temperature variability on *Campylobacter* transmission either, and they did not observe a statistically significant effect with rainfall.

Interestingly, the negative parameter estimate of the variable representing temperature is almost stable throughout the modelling procedure suggesting that higher disease rates per TLA are associated with lower average temperatures. According to the hypothesis of this analysis, this is unsurprising as the pathogen's survival in the environment is advantaged by lower temperatures. However, the final model, which corrects for multicollinear effects, suggests that higher rates of campylobacteriosis occur in TLAs with higher average temperatures and less rainfall. In the New Zealand context, there is some evidence that this might be due to increased faecal material from agricultural activities and more frequent outdoor events and barbeques during the summer months (Eyles et al., 2003). As outlined earlier, both of these activities increase the probability of becoming infected as

exposure to potentially contaminated water, food or animals might be higher in the warmer seasons.

Moreover, it was hypothesised that higher rates of campylobacteriosis appear in TLAs with a high weighted index of interseasonal change. Hearnden et al. (2003) demonstrated the relation between highest summer incidence and greatest inter-seasonal variation for New Zealand at the TLA-level. As the association is stable throughout the modelling process, the results of this analysis support the findings of previous studies investigating the seasonality of campylobacteriosis in New Zealand and other affluent nations (Hearnden et al., 2003; Kovats et al., 2005; Louis et al., 2005; Nylen et al., 2002). However, statistical significance of this relationship is reached only in the first modelling instance. When adjusting for other variables of the category climate/seasonal change, the effect becomes statistically insignificant. Although the level of significance changes, the constant positive direction of the association provides some evidence for the statistical stability of the result.

Landuse

The main land cover class in New Zealand is pastoral (40%) (Khan et al., 2005), and the results of the analysis demonstrate a stable relationship between rural landuse, expressed in livestock density, and the *Campylobacter* rate. The association remains once other factors such as climate and seasonality are controlled for. As the causal connection between livestock density and an increased probability of becoming infected with campylobacteriosis has been described in many national and international studies (e.g. Eyles et al., 2003; Khan et al., 2005; Stanley and Jones, 2003), this finding is consistent with previous research.

Moreover, data from New Zealand's landcover classification (Thompson et al., 2003) were used to assess the probable health risk of agricultural activities (variables representing pastoral landuse). The associations of these data remain unstable, and it could not be shown that particular landuse classes are significant independent factors in explaining the *Campylobacter* rates. Eyles et al. (2002) also describe difficulties in using land cover data to assess the human health risk of campylobacteriosis. As described earlier, the relationship between land cover and the distribution of *Campylobacter* is complex, and other environmental factors including rainfall and presence of livestock have to be taken into account for a comprehensive risk assessment (Eyles et al., 2002). The regression model

"conditions affecting the distribution of the organism" demonstrates that the variables representing seasonality and livestock density only have stable relationships with the disease rate once rainfall and temperature are controlled for.

There is further evidence from New Zealand for the challenge of quantifying the plausible relationship between campylobacteriosis and different types of landuse. McBride et al. (2002) compared the distribution of *Campylobacter* in five areas, each having a predominant landuse type feature: avian, urban, dairy farming, sheep farming, or forestal characteristics. Their findings indicate that the distribution of *Campylobacter* is relatively similar between the different landuse categories. With respect to the results of this thesis, further investigation is necessary to quantify conclusively the relationship between landuse and the distribution of campylobacteriosis.

6.2.2 Exposure

Water

It was hypothesised that higher rates of campylobacteriosis appear in TLAs where the *E.coli* exceedances in drinking water are high because its presence in water is an indicator for faecal contamination. This was, for example, demonstrated by McBride et al. (2002), who found a moderate correlation between *E.coli* and *Campylobacter* for different sampling sites across New Zealand. In turn, this might be related to a greater risk of becoming infected with campylobacteriosis. The inverse relationship observed here, however, suggests that higher rates of campylobacteriosis occur in TLAs with lower levels of *E.coli* exceedances. This might be explained due to conflicting evidence regarding the relationship between indicator organisms, such as *E.coli*, and the actual presence of *Campylobacter*. Frequently, the concentration of *E.coli* is too low to detect faecal contamination of water. Even if the indicator organism is not detected, the water might be contaminated nevertheless (Auckenthaler and Huggenberger, 2003; Hrudey and Hrudey, 2004). Conversely, a high level of *E.coli* exceedances in water is not a conclusive indicator for contamination with *Campylobacter* (Savill et al., 2001). In conclusion, the measurement of *E.coli* exceedances incorporates a large margin of error due to inconsistencies in reporting, sampling practices and analytical methods. Therefore, it is difficult to assess the health risk of *Campylobacter* using this indicator (Khan et al., 2005).

A similar relationship is demonstrated for the association between rainwater and the disease rate, suggesting that lower rates of campylobacteriosis occur in TLAs with a high proportion of people being supplied with rainwater. Compared to national and international findings, this result is surprising. As outlined earlier, rainwater is supposed to have a high risk of being faecally contaminated (e.g. by birds). A recent study investigating the microbiological quality of roof-collected rainwater of private dwellings in New Zealand concludes that:

> *"[a]t least 50% of the roof-collected rainwater samples from private dwellings in New Zealand exceeded the minimal acceptable standards for contamination and 30% of the samples showed evidence of heavy faecal contamination" (Abbott et al., 2005: 16).*

However, it has to be kept in mind that the proxy variable representing the percentage of people with rainwater supply is only an indirect measurement for having a potentially higher risk of becoming infected by contaminated water. The true quality of the rainwater supplies could not be assessed in this thesis. Further evidence for the result being counterintuitive derives from a British study (Said et al., 2003). In England and Wales, 36% of drinking water outbreaks from 1970 - 2000 were associated with private water supplies, although less than 1% of the population is dependent on private water sources. Four main factors are supposed to be major contributors for becoming infected from these supplies: temporary or transient populations without immunity, treatment issues, the presence of animals and heavy rainfall preceding the outbreaks. These factors are also relevant in the New Zealand context.

Moreover, the results of this analysis provide some evidence that higher rates of campylobacteriosis are related to an increased number of drinking water measurements, supposing that there is a greater need for controlling water in some TLAs than in others as the intensity of monitoring reflects action taken by councils and water suppliers (Public Health Surveillance, 2007). Although the relationship is unstable throughout the modelling process, indicating that the effect is not independent, the association shows a positive parameter estimate once all other determinants are controlled for. This relationship, however, does not reach statistical significance.

Due to these arguments, the analysis of the water related variables does not provide enough evidence for the conclusion that water is or is not a significant

independent contributor within the cycle of infection. National and international findings, however, have demonstrated that the percentage of cases resulting from contaminated water supplies might be considerably underestimated (e.g. Abbott et al., 2005; Kapperud et al., 2003; Nygård et al., 2004; Wilson, 2005). In terms of water, further national and international investigations are required in order to quantify its independent effect within the epidemiology and ecology of campylobacteriosis.

Rurality and occupation

All of the models found a consistent relationship between TLAs with a high proportion of people living in highly remote areas and the disease rate. This suggests that the probability of becoming infected with campylobacteriosis is particularly high in rural areas. This result is in line with many national and international studies (e.g. Potter et al., 2003; Sopwith et al., 2006; Wilson, 2005) which particularly emphasise the high risk of falling ill due to direct exposure to potentially contaminated farm environments (e.g. direct contact with infected animals, and water contaminated by agricultural run-off). Therefore, the results of this thesis demonstrate a plausible relationship between factors generally classified into the category “underlying epidemiology” and the *Campylobacter* rate.

The literature examining the role of occupational exposure to campylobacteriosis, however, is inconclusive. Adak et al. (1995), for example, showed that occupational exposure to raw meat is a significant independent risk factor for falling ill; occupational contact with livestock or their faeces, however, was a significant protective factor. Moreover, two studies from the USA demonstrated the relation between *Campylobacter* and poultry processing. Keener et al. (2004) conclude that reducing the contamination of raw poultry on farms and in poultry slaughter plants would have a large impact in reducing the incidence of campylobacteriosis. Potter et al. (2003) found statistical evidence for persons engaged in poultry husbandry having increased odds of infection with *Campylobacter*.

The results of this thesis demonstrated that there was not a consistent occupational effect, and the variables that were intended to represent a potentially higher risk of becoming infected from working in the meat and dairy related industries or direct contact with livestock were not significant in any of the models. Moreover, due to multicollinearity, both of these variables were excluded from the final model.

For New Zealand, Eberhart-Phillips et al. (1997), however, showed that occupational contact with bovine carcasses was strongly associated with campylobacteriosis. Thus, the result of this thesis is not in line with other findings concerning the role of occupational effects in New Zealand, and further investigation is necessary to clarify its relationship with the disease rate.

Food

As demonstrated earlier, campylobacteriosis is primarily described as a foodborne disease, and there is national and international evidence for an increased risk of becoming infected from eating meat prepared in restaurants (Adak et al., 2005; Eberhart-Phillips et al., 1997; Friedman et al., 2004; Gillespie et al., 2003; Kapperud et al., 2003). A proxy was developed in order to assess the potentially greater risk for infection in TLAs with a high restaurant density where the probability of eating out was assumed to be more frequent than in TLAs with a lower restaurant density. Even though this relationship remains statistically insignificant in all models, the association is stable within the modelling procedure. However, the relationship is counterintuitive as the parameter estimate suggests that campylobacteriosis was higher in TLAs with a lower restaurant density. This effect might be a result of applying a proxy variable as a measurement for potentially contaminated foods. The actual quality of food consumed in restaurants could not be directly investigated within the scope of the thesis. However, if the result is a true effect, it is not in line with previous research. Moreover, fast food and barbecued food have been linked to poor hygiene and an increased risk of becoming infected with campylobacteriosis (Eberhart-Phillips et al., 1997; Evans et al., 2003; Gillespie et al., 2003). This effect was not found in this analysis as the association of the developed proxy remains unstable throughout the modelling procedure and shows a negative parameter estimate once all other variables are controlled for.

Further, retail meat products, especially fresh chicken, have been linked to an increased risk of falling ill (Baker et al., 2006b; Michaud et al., 2004; Nielsen et al., 2006; TeckLok et al., 2006). The final model of this thesis shows a positive association between supermarket density and the disease rate when all other variables are controlled for. However, the relationship does not remain constant during the modelling process, nor is the association statistically significant. Even if this is a true outcome, the effect is not independently associated with the *Campylobacter* rate.

The proxy developed to assess an increased probability of becoming infected from consuming fresh food, however, shows an expected and stable association throughout the whole modelling procedure. Moreover, the association is very close to reaching statistical significance in the last modelling step. This result suggests that an increased probability of consuming fresh food might have an independent effect on *Campylobacter* incidence, and this finding is in line with previous research in New Zealand (Eberhart-Phillips et al., 1997). Internationally, Evans et al. (2003) provide evidence for an increased risk of becoming infected with campylobacteriosis from eating fresh vegetables, and Kapperud et al. (2003) identify the consumption of fresh fruits as a protective factor. Although the consistent relationship demonstrated in this thesis provides some evidence for the stability of the statistical result, it has to be kept in mind that the actual food quality was not examined, and it was not possible to assess the fresh fruit and fresh vegetable quality separately.

6.2.3 Characteristics of the affected population and surveillance

Demography

None of the variables representing gender or ethnicity show the expected associations with the *Campylobacter* rate once other factors are controlled for, and the relationships do not reach the level of statistical significance. Further, it was hypothesised that higher rates of campylobacteriosis appear in TLAs with a high proportion of children under 15 as children are particularly at risk of becoming infected (Baker et al., 2006a). However, the results of the regression analysis show a stable inverse association throughout the modelling process, suggesting that higher rates of campylobacteriosis are associated with a lower proportion of children under 15 per TLA. Thus, the results of this analysis are not in line with national and international findings (Mitchell et al., 2002; Morgan, 2002) as the expected patterns are not reflected in the notification data analysed here.

Moreover, the variable representing the proportion of children per TLA is significantly correlated with both of the variables representing the percentage of Europeans per TLA and deprivation. It is likely that the application of crude rates in combination with multicollinear effects causes a distortion of the results.

Conversely, the association between TLAs having a high proportion of people aged 25 to 44 and an increased rate of campylobacteriosis is in line with national and international findings (Baker et al., 2006a; Mitchell et al., 2002; Morgan, 2002).

All of the models found a consistent relationship between this age group and the disease rate. Moreover, the association remains constant once all other variables are controlled for, and the relationship reaches the level of significance in three of the five regression steps, most notably in the final model.

Deprivation

There has been little evidence of a distinct association between social deprivation and campylobacteriosis (Hales et al., 2003; National Health Committee, 2002). For this study, it was hypothesised that higher rates of campylobacteriosis would appear in TLAs where the deprivation index is high. An inverse relationship, however, was observed and suggests that TLAs with a low deprivation index are more likely to experience higher disease rates.

This might, for example, be explained by potential differences in consumption patterns between less deprived areas and more deprived areas. Generally, it is estimated that nutrition related risk factors such as inadequate vegetable and fruit intake contribute to about 40% of deaths annually in New Zealand (Stefanogiannis et al., 2005). Paraphrasing Winkler et al. (2006), the environment (e.g. existence of fresh fruit and vegetable outlets) and individual factors (e.g. access to private transport) impact on dietary behaviour which is, in turn, influenced by socioeconomic characteristics (e.g. income).

In terms of campylobacteriosis, Evans et al. (2003) and Eberhart-Phillips et al. (1997) identified the consumption of fresh vegetables, drinking bottled water and eating out as independent risk factors of becoming infected. These patterns of consumption tend to be associated with a lower deprivation index. For example, Shohaimi et al. (2004) investigated the association between deprivation and the consumption of fresh foods in Norfolk, UK. Their results show that living in a deprived area independently predicts lower consumption of fresh fruits and vegetables. Similar results are, for example, published for studies from Australia (Giskes et al., 2002), Sweden (Lindstrom et al., 2001) and the USA (Deshmukh-Taskar et al., 2007).

Linking the identified inverse association between deprivation and the disease rate to the evaluated positive association for the variable fresh food outlet density, this thesis provides some statistical evidence for the interconnection between deprivation, fresh food and the disease rate. This is particularly interesting in the New Zealand context where prior research has not identified a deprivation gradient in

unhealthy lifestyles in terms of tobacco use, the quantity and pattern of alcohol consumption, levels of physical activity or the intake of fruits and vegetables. (Tobias et al., 2007). In conclusion, the result of this thesis is contrary to most other health outcomes which primarily demonstrate a positive association between deprivation and morbidity/mortality (Pearce and Dorling, 2006; Tobias and Cheung, 2003). However, as mentioned earlier, the actual food quality was not assessed in this research; thus, further investigation is necessary to validate this result.

Artefacts

Generally, under-notification is difficult to assess, and recent studies apply, for example, capture-recapture techniques for the evaluation of undercount (Hoque et al., 2005). In terms of campylobacteriosis, under-notification has not been quantified for the whole of New Zealand. However, a study from Auckland investigated under-notification of giardiasis, another food and waterborne gastrointestinal disease, and the authors conclude that only 49% of the actual cases were notified (Hoque et al., 2005). This indicates that the notified *Campylobacter* cases might only capture a fraction of the true prevalence of the disease.

The variable representing GP density was developed as a proxy for potential under-notification. It was hypothesised that higher rates of campylobacteriosis appear in TLAs where there is better access to health care, represented by the GP density. The results of the regression analysis show a stable negative, but non-significant relationship between GP density and the disease rate. It is concluded that this proxy reflects rurality rather than being a measurement for potential under-notification as the number of GPs per population declines with decreasing population density (Ansari et al., 2003; Wilkinson and Flegg, 2001). This assumption is supported by the finding that highly remote areas appear to have an increased risk of higher disease rates. However, the actual access to health care and the potential impact of under-notification could not be measured within the scope of the thesis.

6.2.4 How do theory and observation fit?

The analysis of the residuals provides the possibility of identifying those TLAs which have a large positive or a large negative deviation between the observed and the expected values. Mapping these discrepancies highlights areas where the rates of campylobacteriosis are higher or lower than expected, once all identified determinants of the disease are controlled for (Clark, 1967; Luke et al., 2000). Here, the residual values were calculated for the final model, which integrates all of the identified determinants of campylobacteriosis and corrects for the issue of multicollinearity.

The identification of areas where the residuals are higher than expected is useful information for planning public health campaigns because territories are highlighted where the model does not explain the occurrence of high disease rates accurately (Shaw and Wheeler, 1985). With respect to the positive residuals, the model reflects the hypotheses of the analysis relatively accurately in 30 out of 38 TLAs (79%). For the other eight TLAs, the model explains the variation in the disease rate less precisely; those TLAs are Thames-Coromandel District, Wellington City, Timaru District, South Wairarapa District, Waikato District, Clutha District, Napier City and Hastings District. The variation in the disease rate appears to be associated with variables other than those identified and included in the model (e.g. impact of foreign travel, cross- contamination from sewage).

With respect to the negative residuals, the model explains the variation in the disease rate relatively accurately for 24 out of 34 TLAs (71%). For the other ten TLAs, the model describes the appearance of the disease less precisely; those TLAs are Tararua District, Carterton District, Marlborough District, Waitakere City, Central Otago District, Manawatu District, Tasman District, Palmerston North City, Masterton District and Nelson City. It would be particularly interesting to understand why the disease rates in these ten TLAs are distinctively lower than expected as this might identify protective factors. This could provide useful information for reducing high rates of campylobacteriosis in other areas.

6.3 Limitations

Applying an ecological study design, this research uses routinely collected data to investigate plausible explanations for the spatial distribution of campylobacteriosis. The evaluation of the results is confronted with a number of challenges, mainly due to notification issues, data availability, the interpretation of proxy information and general limits of an ecological study design (e.g. MAUP, ecological fallacy).

Notification issues

In terms of generating hypotheses on the causation of diseases and monitoring disease trends, population-based morbidity data are essential for surveillance planning and health protection (Hoque et al., 2005). However, there are several reasons why the results of this thesis might be biased. First, reporting and sampling practices might vary between different regions; therefore, the true prevalence in a certain region might not be reflected by the notification data. Second, people might not seek medical attention for various reasons (e.g. unequal access to health care facilities, different sense of sickness); hence, their illness is not notified which in turn results in an undercount of cases for a particular region. Third, the disease is not specifically diagnosed as campylobacteriosis, but more generally as gastrointestinal infection. Fourth, people who are permanently exposed to, for example, a contaminated water source are likely to experience an asymptomatic form of campylobacteriosis which is not notified either (Abbott et al., 2005).

Within the scope of this thesis, it was not possible to quantify these issues; thus, it is difficult to differentiate precisely between the following questions:

- Is the risk of becoming infected really higher/lower in one TLA compared to another?
- Is the notification of cases higher/lower in one TLA compared to another?

Some of the patterns observed here, however, show statistically consistent and contextually plausible associations indicating that the effects might be true. Admittedly, further research is necessary to validate these results, and it is crucial to consider and assess the potential impact of differences in notification practices.

Data availability

As mentioned earlier, this investigation is based on the interpretation of crude rates. It was outlined that age, ethnicity and gender are plausible factors assumed to be affecting the distribution of campylobacteriosis. Even though the modelling process incorporates variables representing demographic characteristics of the population particularly at risk of becoming infected, it is anticipated that the application of standardised data (age, ethnicity, gender) would enhance the model's validity.

In terms of modelling climatic factors assumed to be affecting the distribution of campylobacteriosis, it is challenging to assess this impact at the TLA-level because the spatial units are relatively large. According to New Zealand's topography, climatic conditions vary distinctively within narrow confines, and it is difficult to characterise extreme weather conditions such as heavy rainfall over a large area (Kovats et al., 2005; Sandberg et al., 2006). However, as demonstrated in various case studies (e.g. Savill et al., 2003; Savill et al., 2001), this does not mean that the hypothesised association does not exist or is not important.

Further, it was not possible to consider the plausible transmission route of becoming infected due to recreational water contact. Although NZPHO (2006) provides two proxies related to recreational waterborne diseases[25], the data are either extremely incomplete or do not differentiate between particular waterborne infections. Moreover, a study investigating the hazards of healthy living (Evans et al., 2003) found evidence for a positive association between infection with *Campylobacter* and the consumption of cold tap water and bottled water. Surveillance data regarding these transmission routes are not available for New Zealand. Another plausible source of drinking water contamination is the growing pressure on water quality from wastewater, especially sewage pollutants such as rotting organic matter, excess nutrients and suspended solids. Due to cross-contamination (e.g. leaking water-pipes) the drinking water can be contaminated (Jones, 2001; National Health Committee, 2002). As sewerage data were not available for this thesis, these transmission routes could not be included in this analysis.

[25] Environmental health indicators concerning recreational water (NZPHO, 2006): recreational water quality by TLA in % of marine samples exceeding standard and recreational waterborne disease rate by TLA.

With respect to the food related determinants of campylobacteriosis, the consumption of contaminated poultry is supposed to be a main transmission route for becoming infected, and this is of particular importance in the New Zealand context (Baker et al., 2006b). Although this study includes some variables representing a potential risk of becoming infected from contaminated foods, it was not possible to account specifically for the increased consumption of poultry over the last decades. Poultry consumption in New Zealand has increased from 1kg per person in the 1960s to 34.9 kg in 2002 (MAF, 2003).

The literature also reports that many cases of campylobacteriosis are associated with foreign travel, mainly to countries in Southern Europe, India and East Africa (Ekdahl and Andersson, 2004; Norström et al., 2006). This relationship is not quantified in the New Zealand context; thus, it was not possible to exclude those cases from the analysis which have a non-New Zealand aetiology.

Finally, it was not possible to separate sporadic cases of campylobacteriosis from outbreak related cases which are characterised by two different aetiologies (Gillespie et al., 2003; Potter et al., 2003). This issue is also closely related to suspected differences in the ecology of rural and urban exposures (Skelly et al., 2002). In particular, water associated outbreaks which could be traced back to cross-contamination of the public water supply and foodborne outbreaks are related to urban areas rather than rural areas because of different lifestyles and consumption patterns (Nygård et al., 2004). However, outbreak related data were not available for the TLA-level. Moreover, the scope of statistical analyses would be limited as few disease outbreaks and cases have been linked conclusively to one particular source (ESR, 2001 - 2005). This has to be considered when interpreting statistical results: the fewer the observations, the less reliable the statistical results.

Evidence of multicollinearity and heteroscedasticity

Five proxies were excluded from the last step of the hierarchical regression procedure in order to correct for multicollinear effects which occur when two or more of the independent variables are correlated (Norušis, 2004). Thus, the final model does not account for the assumed impact of the variables representing primarily pastoral landuse, stock density per population, proportion of people in dairy and meat related industries, and both of the demographic variables age and ethnicity. It is expected that the inclusion of other proxies representing these plausible

determinants of campylobacteriosis would enhance the model's degree of explanation.

Although it is possible to compare the magnitudes of *beta* coefficients for the purpose of deducing the relative importance of variables within the regression model, it has to be kept in mind that *beta* coefficients are affected by the variability of the analysed association. If the assumption of homoscedasticity is seriously violated, the interpretation of the *beta* coefficients can be biased when there is a high variability in the associations between the dependent and the independent variables (Macfie and Nufrio, 2006; Pedhazur, 1997). With respect to the analysis of the residuals (figure 17c, scatterplot), there is some evidence that the findings might be biased due to heteroscedasticity. However, this does not mean that the observed effects are insignificant or even dispensable. Even small ecological effects may have significant public health implications (Blakely and Woodward, 2000; Rothman and Greenland, 2005).

Modifiable area unit problem

A well described issue in spatial analysis is the modifiable area unit problem concerning scale and aggregation. The results of this analysis might be biased due to a greater instability of rates in TLAs with small populations where the likelihood of computing extreme values rises (e.g. Huruni district, Banks Peninsula district) (Hertz-Picciotto, 1998). Smoothing techniques, for example, can de-emphasise such effects. However, these techniques are most suitable for smaller geographical areas such as CAUs or meshblocks (Rezaeian et al., 2007).

Ideally this analysis, which is conducted on the TLA-level, could have been executed on the CAU or meshblock level in order to compare the stability of the effects on different spatial scales. Duplicating results on different modelling levels would confirm the stability of the statistical results; contradictory findings, however, would indicate the opposite. Rican et al. (1999), for example, investigated geographic differences of bronchopulmonary cancer mortality in France on different spatial scales. The results show different interactions between environmental and social factors on each scale. Thus, the authors conclude that the validity of a regional scale to study health geographical distributions is questionable. Observing stable results on different spatial scales as well as throughout a hierarchical modelling process would confirm plausible relationships between the distribution of a disease

and its assumed determinants. However, most of the environmental data used for this analysis (e.g. climate data, stock units, landuse data) are not available for the CAU or meshblock level.

Ecological fallacy

The ecological fallacy refers to the problem of inferring causality from the observed associations (Rezaeian et al., 2007). *Campylobacter* rates on the TLA-level show the average morbidity experience for people living in this particular area for the period 1997 - 2005. The identified plausible determinants of campylobacteriosis also represent average values; additionally, most of these variables were only collected for a particular year. Although it was possible to calculate the theoretical risk of becoming infected with campylobacteriosis per TLA and evaluate how far this might reflect the observed *Campylobacter* rates, association does not imply causation. For example, the final model explains the high rates of campylobacteriosis in the Hurunui district relatively well. However, this does not mean that all individuals living in this district share the same extent of exposure to a particular source of infection. Moreover, the extent of a particular exposure might have changed between 1997-2005; this could not be considered in the scope of this thesis. Additionally, this association might be biased due to the prior mentioned MAUP, as the Hurunui district is one of the less populated TLAs in New Zealand.

6.4 Current public health interventions and implications of research

Although the conclusive verification of a scientific process is theoretically impossible (Hill 1965, in: Rothman, 1986), action is required if a situation such as increasing rates of campylobacteriosis is classified as hazardous for public health. Taking action might even be successful, although the reasons for a particular implementation could be refuted by subsequent studies.[26] Hence, it might be reasonable to put preventive measures into action despite the knowledge that uncertainties influence the decision process. Public policy is based on the

[26] For example, research by the German chemist and hygienist Max Joseph von Pettenkofer (1818 - 1901) was concerned with the epidemiology of cholera. His findings resulted in the reconstruction of the drinking water and sewage system in Munich, and subsequent cholera epidemics could be controlled more effectively. He associated the distribution of this disease with contaminated soil which in turn contaminated the water supply; subsequent research, however, identified water as the determining transmission route for the pathogen *vibrio cholerae*.

interpretation of the best available information; it does not require absolute scientific certainty (Phillips and Goodman, 2004; Schonwalder and Olden, 2003).

6.4.1 The "chicken debate" and current public health interventions

One of the major possible public health implications discussed in the context of reducing high rates of campylobacteriosis is the debate regarding the banning of fresh poultry from sale. On the one hand, various studies argue that there is strong epidemiological and laboratory evidence that fresh chicken is the dominant source of human infection with *Campylobacter* (Baker et al., 2006b). For New Zealand, there are two major case-control studies confirming high risk for becoming infected with campylobacteriosis from fresh chicken (Eberhart-Phillips et al., 1997; Ikram et al., 1994). Paraphrasing Baker et al. (2006b) and Reiersen et al. (2002), the rising disease rates have coincided with a significant increase in poultry consumption, and there is international evidence for the successful decrease in high rates due to banning fresh poultry from sale. In Belgium, for example, fresh poultry was banned from sale for four weeks during 1999 as this food was possibly contaminated by dioxin. The disease rate dropped significantly and rose again when poultry was reintroduced. Further evidence derives from experience with frozen chicken in Iceland; since 2000, all known *Campylobacter*-positive broiler flocks are frozen before they go to retail. In combination with extensive consumer education, these implementations have also resulted in a significant reduction of the domestic *Campylobacter* rate.

On the other hand, a major argument of a recent study (Nelson and Harris, 2006) deals with the serotype analysis of the *Campylobacter* isolates, which is ambiguous and inconclusive regarding chicken meat itself as the major source of infection. This was earlier suggested by Cox (2004), who pointed out that the overlap between human and poultry serotypes might be due to the same source of infection, and should not be interpreted as the conclusive factor for poultry being the most significant contributor to food related *Campylobacter* infections. Moreover, Nelson and Harris (2006) emphasise the disjunction between seasonal patterns of human and chicken infection, as the seasonal colonisation pattern in live chickens follows the seasonal increase in human infections. They conclude that chicken consumption is a significant risk factor, but regarding poultry as the conclusive major source of infection is forejudging, and it is precipitate to neglect the consideration of other sources of the infection such as other foods or water.

At the moment, New Zealand's health authorities and policy makers do not regulate the retail of fresh chicken. With respect to the prior outlined debate, the current policy approach supports the latter argument that there is not enough evidence for taking such drastic action as banning fresh chicken from sale, which would also have severe economic consequences for the poultry industry.[27] However, as described above there is debate about this possible intervention.

Already implemented public health interventions were mainly initiated by the Ministry of Health and the Institute of Environmental Science and Research. These interventions comprise, for example, the implementation of *Public Health Risk Management Plan Guides for Drinking Water Supplies*, the *Public Health Grading of Community Drinking - Water Supplies*, food safety programmes, educational preventive initiatives, and the improvement of surveillance activities, e.g. through the provision and enhancement of health and health related data (ESR, 2006a; MoH, 2005 -a; MoH, 2005 -b; MoH, 2005 -c; Weinstein et al., 2000). However, regarding the high rates of campylobacteriosis, these interventions appear to be insufficient.

6.4.2 Implications of research for public health authorities to consider

Surveillance data are a basic requirement for risk analysis and the reduction of high rates of campylobacteriosis, which are key priority areas for the New Zealand Government (2007). In particular, the notifiable disease and environmental health indicator dataset provided by NZPHO (2006) is a useful tool to monitor trends in the development of campylobacteriosis. Furthermore, the dataset includes water related determinants of campylobacteriosis; however, as highlighted by this research, this dataset suffers from deficiencies, and it was particularly difficult to quantify the water-related impact on campylobacteriosis. Further, it would be very useful to include outbreak related surveillance data in this database in order to separate sporadic from outbreak related cases. It is concluded that the TLA-level is a practical spatial unit to identify potential "hotspots" of campylobacteriosis for the whole of New Zealand, which is useful information for planning future research on a smaller spatial scale such as the CAU or meshblock level. In terms of improving the validity of the results at the TLA-level, the enhancement of the quality and availability of data included in this particular dataset would be desirable.

[27] In 2002, the poultry meat industry earned NZ $507 million in retailing sales (MAF, 2003).

As previously stated, the interpretation of proxy information is inherently challenging, especially with respect to the water, rurality and food related contexts investigated here. Moreover, most of these data are maintained by different organisations which possibly apply varying methods of data collection (e.g. ESR, NZPHO, SNZ, MAF). The above mentioned database maintained by NZPHO (2006) aims to integrate various environmental health indicators; however, further harmonisation of different data sources would be helpful for increasing the comparability of data.

Further, the importance of poultry as a major transmission route was described. However, the data concerning poultry production and consumption are extremely fragmentary or not available at all. Reasons for the incompleteness of these data are mainly of a confidential nature. However, what is actually meant by the statement *"data have been suppressed for reasons of confidentiality"* (MAF, 2002b: 1) is not further elaborated. In terms of the high rates of campylobacteriosis it is incomprehensible why only these data are unavailable, whereas all other livestock and domestic animal related data are relatively comprehensive and complete (MAF, 2002a; MAF, 2003). In terms of research on campylobacteriosis, the availability of poultry related data is crucial.

Finally, deprivation was identified as a significant factor in explaining the observed spatial patterns in the notification data. As outlined earlier, only a few studies have paid attention to this relationship, especially in the New Zealand context. This might be related to the situation that studies concerned with the investigation of campylobacteriosis are mainly focusing on the ecological and food-borne dimension of the disease. Particularly from an epidemiological and ecological point of view, this remains important as there are still major gaps in existing knowledge (Devane et al., 2005; WHO, 2000). However, a comprehensive public health approach might want to pay more attention to relatively "unexplored" dimensions of campylobacteriosis.

6.6 Summary

The discussion explained and evaluated the effects of a number of plausible determinants of campylobacteriosis in the context of national and international findings. There is statistical evidence that some of the investigated relationships such as seasonality, rurality, age and deprivation appear to have independent effects on the distribution of the disease rate. Other hypotheses, such as the assumed effects of rainfall and occupational contact to potentially contaminated animals or meat, could not be confirmed by this research.

Theoretical context provides the frame in which statistical probability has to be evaluated. However, cause and effect relationships cannot be proved statistically as statistical significance is at best tentative and a subjective process dependent on current knowledge (Rothman and Greenland, 2005). Further, the findings might be biased due to the reasons listed below:

- notification artefacts (e.g. in particular population groups, infections might be more likely to be notified than in others, or there might be inconsistencies in reporting),
- the usage of proxy variables including the issue of multicollinear associations,
- general limits of an ecological study design (MAUP, ecological fallacy).

The discussed limitations of the study reflect the challenge of modelling the complex transmission routes of campylobacteriosis, especially when applying a comprehensive modelling approach dealing with more than one particular source of infection. The implications of this research are primarily concerned with the enhancement of the quality and availability of current data sources. Moreover, this study found evidence for currently relatively unexplored dimensions of campylobacteriosis, such as the statistical association between food, deprivation and the disease rate.

7. Conclusion

This study examined the geographical distribution of campylobacteriosis in New Zealand and gained insight into the relative importance of plausible determinants assumed to be affecting the observed spatial patterns. This was achieved by examining all rates of campylobacteriosis together for the period 1997 - 2005. Moreover, plausible determinants of the disease which may be particularly important in the New Zealand context were identified. Hierarchical regression was applied in order to quantify the magnitude of effects in their complex interrelation. Although the study design contains some challenges, it was possible to analyse a complex environmental health relationship for the whole of New Zealand. The final chapter summarises key findings of the analysis and illustrates the possibilities of potential future research.

7.1 Key findings of the analysis

The analysed notification data (1997 - 2005) show a large geographical variation, and there was an increase of rates for nearly all TLAs across New Zealand. The statistical analysis confirms evidence for local and global clustering; higher rates, however, primarily appear in the South Island. The observed spatial trends are in accordance with national and international findings focusing on campylobacteriosis in the context of affluent nations.

Generally, the modelling approach partially explains the variation in campylobacteriosis and confirms statistical significance for some of the independent variables throughout the modelling process. The final regression model is a summary of the data, corrects for multicollinearity and describes how all variables are linked. The validity of the results, however, has to be confirmed applying a study design particularly dealing with the issues described as MAUP, ecological fallacy and notification artefacts.

In particular, the hierarchical regression provides statistical evidence for the significant contribution of some of the variables classified in the category “characteristics of affected population/surveillance factors”. Most notably, the results show a consistent and statistically significant negative association between deprivation and the disease rate throughout the entire modelling procedure. This

result is of particular relevance as it is contrary to most other health outcomes describing higher morbidity and mortality rates for more deprived areas.

Although not statistically significant, some of the variables classified in the category "underlying epidemiology" show consistent associations with the *Campylobacter* rate. In particular, seasonality, stock density, the proportion of people living in highly remote areas and the fresh food outlet density show expected and consistent associations once all other identified determinants of campylobacteriosis are controlled for.

Relating the identified association between fresh food outlet density to that of deprivation and the *Campylobacter* rate links the categories "underlying epidemiology" and "characteristics of affected population/surveillance factors" in explaining the observed disease rates. This result is particularly relevant as prior research has paid little attention to the associations between deprivation, fresh food and campylobacteriosis.

The unexplained variation in campylobacteriosis reflects the challenge of taking into account a wide spectrum of transmission routes and of quantifying the relative contribution of plausible determinants of the disease. The analysis of the final model's residuals identifies hotspots of interest for future research.

7.2 Future research

This research investigated plausible environmental and sociodemographic determinants of campylobacteriosis and their complex interrelations in order to further current knowledge about the ecology and epidemiology of campylobacteriosis. With respect to the prior discussed limitations and the derived implications of the study's results, it is anticipated that future research in the areas described below will substantially improve current understandings of this disease.

7.2.1 Further exploration of determinants of campylobacteriosis

Although the final model of the hierarchical regression explains the observed *Campylobacter* rates in many TLAs relatively well, the analysis of the residuals highlights areas where the model fails to explain the observed disease rate. Generally, future research could concentrate on identifying unspecified determinants of campylobacteriosis and on minimising the multicollinear effects detected here.

Conversely, it might be useful to further investigate areas where the residuals are lower than expected in order to identify and understand protective factors potentially limiting the increase in rates in some areas. Moreover, it would be interesting to quantify the magnitude of determinants of campylobacteriosis separately for urban and rural areas, and outbreaks and sporadic cases respectively. Finally, it is crucial to quantify the impact of notification artefacts.

In particular, it would be interesting to investigate further the possibly underestimated role of water in the circle of infection with *Campylobacter*. Although the associations of the water related variables investigated here are inconclusive, national and international case studies describe the significance of water as a transmission route for campylobacteriosis (Nygård et al., 2004; Savill et al., 2001; Weinstein et al., 2000). However, the different dimensions of this relationship have not been quantified on a broader scale (e.g. impact of recreational water contact, consumption of tap or bottled water, cross-contamination of drinking-water by sewage). An integrated analysis of data provided by the national databases WINZ, CoSINZ and SWSS[28] in a future research project could improve knowledge about the role of water as a determinant of campylobacteriosis. In turn, this would provide new evidence which could be included in a comprehensive modelling approach considering more than one particular source of infection.

Further, this research found statistical evidence for a plausible link between a high fresh food outlet density, less deprived TLAs and an increased disease rate. Generally, there is strong evidence for the relationship between deprivation and health in New Zealand (Pearce and Dorling, 2006). This finding is particularly relevant and worth further investigation as the literature provides only few and inconclusive findings concerning the threefold relationship of deprivation, food and campylobacteriosis. Moreover, none of the studies considered in the New Zealand context control for several plausible environmental and sociodemographic

[28] CoSINZ (Community Sewage Information New Zealand): a national database defining community sewage systems, developed in 2000 to support the need to classify such systems and to determine the upgrade costs of meeting health and environmental concerns (ESR, 2005).

SWSS (Sanitary Works Subsidy Scheme): this database is used nationally by the Ministry of Health to manage applications for sewage system improvement or drinking-water fluoridation subsidies. Each application is based upon the relevant community as defined in CoSINZ or WINZ (ESR, 2005).

determinants on campylobacteriosis in order to explore the particular impact of deprivation on the disease rate.

7.2.2 Strengthening public health awareness

As stated earlier, it is politically and scientifically widely accepted that the risk of becoming infected with campylobacteriosis is neither marginal nor tolerable. Political statements on campylobacteriosis are of increasing importance throughout the media (e.g. the internet), and the scientific literature on campylobacteriosis has been growing since 1968. For example, entering the keyword *"Campylobacter"* into the medical database pubmed resulted in 10,078 matches for 27 March 2007, whereas the keyword "campylobacteriosis" still showed 4,681 matches for the same day. Just referring to the latter, the first and only reference for this subject index is recorded for 1968. Ten years later, there were already 41 entries, and in 2006 the keyword resulted in 172 matches. Although the database provides only nine matches for the keywords "campylobacteriosis, New Zealand" for 2006, various articles, political statements and other media releases have been published dealing with this subject. For instance, a Google search on 27 March 2007 resulted in 21,800 hits for the same keywords, just considering pages from New Zealand.

The prior examples emphasise the political and scientific dimension of this health risk, but the literature dealing with the public risk perception of the subject matter is marginal. However, improving the public's perception and awareness of the high risk of becoming infected with campylobacteriosis is crucial in terms of an effective preventive approach. Quoting Arthur et al. (2002: 7),

> *"[t]he perceptions of consumers and those of the scientific community differ in terms of the relative importance of different hazards that affect food safety."*

A recent study from NZFSA (2005) considers the topic of risk perception of New Zealanders in a quantitative way. About 60% of the respondents declare having "a lot" or "a fair amount" of interest in food safety issues. Generally, females and Maori declare stronger interest in problems related to food safety than males or other ethnicities. Becoming infected from contaminated chicken or shellfish is of particular concern, whereas cheese, canned food, fresh fruit and vegetables are of least concern. With respect to the thesis' findings, which suggest an association almost statistically significant for the relationship between the *Campylobacter* rate and the

fresh food outlet density, the latter is of particular interest and worth further investigation. Furthermore, most of the study's respondents are concerned about food safety issues regarding salmonella, listeria[29], the use of pesticides to grow food, and antibiotics in meat. Interestingly, the rare disease of listeria is noted in this study, whereas campylobacteriosis is not mentioned at all.

Public awareness results from noticing and evaluating available information on campylobacteriosis which is closely related to the presentation of this health risk in the media and one's personal risk perception. People who have suffered from an infection with *Campylobacter*, scientists or medical staff are likely to be more concerned about this health risk. Different courses of action are conceivable: for instance, people might deliberately eat less poultry or try avoiding cross-contamination of food in the kitchen.

Broad levels of the population, however, appear to be unaware of this health risk. For example, an American study (Physicians Committee for Responsible Medicine, 2001) investigating public awareness of campylobacteriosis demonstrates 84% of the respondents being unaware of the source of foodborne diseases, including the aetiology of campylobacteriosis. Amongst others, Miles et al. (1999) describe a decline in consumer knowledge about safe food preparation and the *"optimistic bias"* (Miles et al., 1999: 754) as reasons for increasing rates of campylobacteriosis. Paraphrasing the authors, the *"optimistic bias"* refers to the issue that information about the risk of becoming infected is usually generalised. This allows the conclusion that everybody else, but oneself, is affected by a health risk such as campylobacteriosis.

Future research concentrating on the issue of public health awareness would contribute to a more effective risk management because

> *"[i]t has been seen that people are far less knowledgeable about Campylobacter than Salmonella or E. coli, yet Campylobacter is one of the most common causes of food poisoning. […]. Effective risk communication about microbiological food risks is essential if people are to change behaviours associated, [for example], with poor food hygiene practice" (Miles et al., 1999).*

[29] Listeria is a rare, but lethal foodborne infection. The patient is seriously ill with a high fever, headache and confusion. The disease has various complications and a mortality rate of 25% (Todar, 2003).

The comprehension of environmental health relationships is the first step in understanding health risks such as campylobacteriosis. An approach considering questions such as “Is the risk small? Do we know little about the risk? How do we communicate what we know?” (Fülgraff, 2001: 304, 308) complements a primarily epidemiologic and ecological perspective.

A recent press release from the New Zealand Government (2007) indicates that future research will consider the latter aspects to a greater extent. In 2006/2007, the Government provided $2.4 million to fund 10 new research projects; one of them deals with the investigation of *Campylobacter* in food and the environment, including water as a transmission route of the pathogen and a chapter about risk communication. The research will be overseen by NZFSA and the Ministry for the Environment and carried out by researchers from Massey University, the National Institute of Water and Atmospheric Research and ESR. Moreover, this research is closely related to the NZFSA *Campylobacter* risk management strategy 2006 - 2007 (NZFSA, 2006). The project is seen as “*an excellent opportunity to take a holistic approach to the issue”* (NZFSA, 2007: 1).

7.3 Concluding statement

Especially in the New Zealand context, prior research has neglected to explore the spatial dimension of this disease. Thus, the results of this thesis add to the current discussion on campylobacteriosis as they provide further evidence for the role of place in explaining the distribution of a particular disease. As a result, health authorities should perhaps consider geographically focused projects with a transdisciplinary study approach. In addition, future work needs to integrate relatively unexplored determinants of campylobacteriosis such as the role of deprivation or the public’s perception of the health risk of campylobacteriosis. In summary, it is anticipated that public health interventions and future research in the areas listed below will considerably improve our understanding of campylobacteriosis and contribute to protect public health and cut the costs involved in reducing rates and preventing infections:

- enhancement of the quality and availability of surveillance data;
- improvement of knowledge about the complex transmission routes of campylobacteriosis and conclusive quantification of the relative contribution of each identified source to the overall impact of the disease;

- improvement of knowledge concerning the health risk perception of campylobacteriosis;
- unification of existing knowledge and development of a convincing course of action supported by the major actors concerned.

Campylobacteriosis is a complex, persistent, increasing and expensive public health issue. In order to reduce the high incidence rates of this disease, health authorities need to consider individual determinants (e.g. age, sex, ethnicity, socio-economic position) and area-level characteristics (quality of drinking water) (Bentley and Kavenagh, 2007). Providing safe food and water is an essential requirement for reducing the risk of becoming infected, and water and food related issues are primarily overseen and regulated by the New Zealand Government (MoH, 2005 -b; NZFSA, 2006). Moreover, it is increasingly important to improve the public's awareness and knowledge about this health risk in order to reduce high disease rates in the long term. In conclusion, as long as there is no collective and convincing course of action supported by the scientific community, politicians, the food and water industry, and other main actors contributing to public health, reducing high rates of campylobacteriosis and its financial burden remains particularly challenging.

8. Literature

Aamodt, G.; Samuelsen, S. O. & Skrondal, A. (2006): A simulation study of three methods for detecting disease clusters. *International Journal of Health Geographics 5*, no. 15, pp. 1 - 11.

Abbott, S. E.; Douwes, J. & Caughley, B. P. (2005): A survey of the microbiological quality of roof-collected rainwater of private dwellings in New Zealand. Retrieved: 28/04/2007, from: http://www.water2006.com/manuscripts_06/index.htm.

Adak, G. K.; Cowden, J. M.; Nicholas, S. & Evans, H. S. (1995): The Public Health Laboratory Service national case-control study of primary indigenous sporadic cases of campylobacter infection. *Epidemiolgy and Infection 115*, no. 1, pp. 15 - 22.

Adak, G. K.; Meakins, S. M.; Yip, H.; Lopman, B. A. & O'Brien, S. J. (2005): Disease risks from foods, England and Wales, 1996-2000. *Emerging Infectious Diseases 11*, no. 3, pp. 365 - 372.

Ahlbohm, A. & Norell, S. (1991): *Einführung in die moderne Epidemiologie.* MMV Medizinverlag, München.

Allos, B. M. (2001): Campylobacter jejuni Infections: update on emerging issues and trends. *Clin Infect Dis 32*, no. 8, pp. 1201 - 1206.

Altekruse, S. F.; Stern, N. J.; Fields, P. I. & Swerdlow, D. L. (1999): Campylobacter jejuni - An emerging foodborne pathogen. *Emerging Infectious Diseases 5*, no. 1, pp. 28 - 35.

Anonymus (2002): Surveillance übertragbarer Krankheiten in der Perspektive des Infektionsschutzgesetzes. *Bundesgesundheitsbl-Gesundheitsforsch-Gesundheitsschutz 11*, no. 3, pp. 839 - 843.

Ansari, Z.; Barbetti, T.; Carson, N. J.; Auckland, M. J. & Cicuttini, F. (2003): The Victorian ambulatory care sensitive conditions study: rural and urban perspectives. *Social and Preventive Medicine 48*, no. 1, pp. 33 - 43.

Arthur, J.; Jensen, J. & Spence, H. (2002): Microbiological Foodborne Disease In New Zealand. Current Issues, Problems And Solutions. *Journal of the New Zealand Dietetic Association 56*, no. 1, pp. 7 - 13.

Auckenthaler, A. & Huggenberger, P. (Eds.) (2003): *Pathogene Mikroorganismen im Grund- und Trinkwasser. Transport - Nachweismethoden - Wassermanagement,* Basel, Birkhäuser Verlag AG.

Baker, M. (2006): *correspondence via email, Christchurch, New Zealand. Date: 27/09/2006.*

Baker, M.; Ball, A.; Devane, M.; Garrett, N.; Gilpin, B.; Hudson, A.; Klena, J.; Nicol, C.; Savill, M.; Scholes, P. & Williams, D. (2002): Potential Transmission Routes of Campylobacter from Environment to Humans. Client Report FW0246. Retrieved 23/06/06, from: http://www.moh.govt.nz/moh.nsf/0/D9A729032610C1E6CC256CA5006CFACE.

Baker, M. G.; Sneyd, E. & Wilson, N. A. (2006a): Is the major increase in notified campylobacteriosis in New Zealand real? *Epidemiology and Infection 135*, no. 1, pp. 1 - 8.

Baker, M. G.; Wilson, N.; Ikram, R.; Chambers, S.; Shoemack, P. & Cook, G. (2006b): Regulation of chicken contamination urgently needed to control New Zealand's serious campylobacteriosis epidemic. *New Zealand Medical Journal 119*, no. 1243, pp. 1 - 8.

Bartels, C. J. & Van Beurden, A. U. C. J. (1998): Using geographic and cartographic principles for environmental assessment and risk mapping. *Journal of Hazardous Materials 61*, no. 1, pp. 115 - 124.

Beaglehole, R.; Bonita, R. & Kjellström, T. (2000): *Basic epidemiology.* World Health Organization, Geneva.

Bentley, R. & Kavenagh, A. (2007): Interpreting contextual effects in mulitlevel models of health. *Australasian Epidemiologist 14*, no. 1, pp. 4 - 5.

Blakely, T.; Atkinson, J.; Ivory, V.; Collings, S.; Wilton, J. & Howden-Chapman, P. (2006): No association of neighbourhood volunteerism with mortality in New Zealand: a national multilevel cohort study. *International Journal of Epidemiology 35*, no. 4, pp. 981 - 989.

Blakely, T. A. & Woodward, A. J. (2000): Ecological effects in multi-level studies. *Journal of Epidemiolgy and Community Health 54*, no. 5, pp. 367 - 374.

Brieseman, M. A. (1987): Town water supply as the cause of an outbreak of Campylobacter infection. *New Zealand Medical Journal 100*, no. 821, pp. 212 - 213.

Brieseman, M. A. (1990): A further study of the epidemiology of Campylobacter jejuni infections. *New Zealand Medical Journal 103*, no. 889, pp. 207 - 209.

Briggs, D. J. & Wills, J. (1999): Presenting Decision-Makers With Their Choices: Environment Health Indicators For NEHAPs. In: Briggs, D. J.; Stern, R. & Tinker, T. L. (Eds.): *Environmental Health for All. Risk Assessment and Risk Communication for National Environmental Health Action Plans. NATO Science Series 2. Environmental Security, Vol. 49.* Kluwer Academic Publishers, Dordrecht, Boston, London, pp. 187 - 201.

Brown, T. & Duncan, C. (2000): London's burning: recovering other geographies of health. *Health and Place 6*, no. 4, pp. 363 - 375.

BSES (2006): Bundesbehörden der Schweizer Eidgenossenschaft (BSES). Faktoren für die Umrechnung des Tierbestandes in Grossvieheinheiten. Retrieved: 03/08/2006, from: http://www.admin.ch/ch/d/sr/910_91/app1.html.

Butzler, J. P. (2004): Campylobacter, from obscurity to celebrity. *Clinical Microbiology and Infection 10*, no. 10, pp. 868 - 876.

Campbell, D.; Hathaway, S. & van der Logt, P. (2006): New Zealand Food Safety Authority's response to the 'flies, fingers, fomites, and food' article on campylobacteriosis. *New Zealand Medical Journal 119*, no. 1241, pp. U2157.

Campbell, K. W. & Gilbert, S. A. (1995): *Poultry Quality Assessment.* Public Health Research Commission, Ministry of Health, Wellington.

Carteris, P. L.; Morris, S. T. & Matthews, P. N. P. (2000): Pasture-based beef production in New Zealand. Retrieved: 10/02/2007, from: http://www.beef.org.nz/statistics/plc991.pdf.

Centers for Disease Control and Prevention (2006): Principles of Epidemiology in Public Health Practice. Retrieved: 01/11/2006, from: http://www.cdc.gov/training/products/ss1000/ss1000-ol.pdf.

Clark, W. (1967): The use of residuals from regression in geographical research. *New Zealand Geographer 23*, no. 3, pp. 64 - 66.

Clarke, K. C.; McLafferty, S. L. & Tempalski, B. J. (1996): On epidemiology and geographic information systems: a review and discussion of future directions. *Emerging Infectious Diseases 2*, no. 2, pp. 85-92.

Coker, A. O.; Isokpehi, R. D.; Thomas, B. N.; Amisu, K. O. & Obi, C. L. (2002): Human campylobacteriosis in developing countries. *Emerging Infectious Diseases 8*, no. 3, pp. 237 - 244.

Cox, L. A. (2004): Letters to the Editor: *Campylobacter* Risk Data Out of Date? *Comprehensive Reviews in Food Science and Food Safety 3*, no. n/a, pp. 125 - 126.

Cramer, D. (2003): *Advanced Quantitative Data Analysis.* Open University Press, Maidenhead, Philadelphia.

Cromley, E. K. & McLafferty, S. L. (2002): *GIS and Public Health.* Guilford Press, New York.

Crushell, E.; Harty, S.; Sharif, F. & Bourke, B. (2004): Enteric campylobacter: purging its secrets? *Pediatric Research 55*, no. 1, pp. 3 - 12.

Curriero, F. C.; Patz, J. A.; Rose, J. B. & Lele, S. (2001): The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948 - 1994. *American Journal of Public Health 91*, no. 8, pp. 1194 - 1199.

Deshmukh-Taskar, P.; Nicklas, T. A.; Yang, S. J. & Berenson, G. S. (2007): Does food group consumption vary by differences in socioeconomic, demographic, and lifestyle factors in young adults? The Bogalusa Heart Study. *Journal of the American Dietetic Association 107*, no. 2, pp. 223 - 234.

Devane, M. L.; Nicol, C.; Ball, A.; Klena, J. D.; Scholes, P.; Hudson, J. A.; Baker, M. G.; Gilpin, B. J.; Garrett, N. & Savill, M. G. (2005): The occurrence of Campylobacter subtypes in environmental reservoirs and potential transmission routes. *Journal of Applied Microbiology 98*, no. 4, pp. 980 - 990.

Duncan, C.; Jones, K. & Moon, G. (1998): Context, composition and heterogeneity: using multilevel models in health research. *Social Science and Medicine 46*, no. 1, pp. 97 - 117.

Duncanson, M.; Russell, N.; Weinstein, P.; Baker, M.; Skelly, C.; Hearnden, M. & Woodward, A. (2000): Rates of notified cryptosporidiosis and quality of drinking water supplies in Aotearoa, New Zealand. *Water Research 34*, no. 15, pp. 3804 - 3812.

Eberhart-Phillips, J.; Walker, N.; Garrett, N.; Bell, D.; Sinclair, D.; Rainger, W. & Bates, M. (1997): Campylobacteriosis in New Zealand: results of a case-control study. *Journal of Epidemiology and Community Health 51*, no. 6, pp. 686 - 691.

Ekdahl, K. & Andersson, Y. (2004): Regional risks and seasonality in travel-associated campylobacteriosis. *BMC Infectious Diseases 4*, no. 1, pp. 54 - 60.

Elliott, P.; Martuzzi, M. & Shaddick, G. (1995): Spatial statistical methods in environmental epidemiology: a critique. *Statistical methods in medical research 4*, no. 2, pp. 137 - 159.

Elliott, P. & Wartenberg, D. (2004): Spatial epidemiology: Current approaches and future challenges. *Environmental Health Perspectives 112*, no. 9, pp. 998 - 1006.

English, D. (1992): Geographical epidemiology and ecological studies. In: Elliott, P.; Cuzick, J.; English, D. & Stern, R. (Eds.): *Geographical and Environmental Epidemiology. Methods for Small-Area Studies.* Oxford University Press, Oxford, New York, Tokio, pp. 3 - 13.

ESR (2001): Campylobacter. Retrieved: 02/08/2006, from: http://www.nzfsa.govt.nz/science/data-sheets/campylobacter.pdf.

ESR (2001 - 2005): Annual Summary of Outbreaks. Retrieved: 21/07/2006, from: http://www.surv.esr.cri.nz/surveillance/annual_outbreak.php.

ESR (2002): Campylobacteriosis occurring at record levels in December, 2001. *New Zealand Public Health Report 9*, no. 1, pp. 5 - 7.

ESR (2005): Water Information Systems. Retrieved: 21/11/2006, from: http://www.esr.cri.nz/competencies/watermanagement/waterinformationsystems.htm.

ESR (2006a): National WINZ Database. Retrieved: 23/08/2006, from: http://www.drinkingwater.co.nz/external/downloads.asp.

ESR (2006b): Notifiable and other diseases in New Zealand. Annual Report 2005. Client Report FW 0621. Retrieved 23/07/2006, from: http://www.surv.esr.cri.nz/surveillance/annual_surveillance.php?we_objectID=1016.

ESR & MoH (2006): Drinking water for New Zealand. Water information New Zealand (WINZ). Retrieved: 17/10/2006, from: http://www.drinkingwater.co.nz/supplies/explainregisterentry.asp.

Evans, M. R.; Ribeiro, C. D. & Salmon, R. L. (2003): Hazards of healthy living: bottled water and salad vegetables as risk factors for Campylobacter infection. *Emerging Infectious Diseases 9*, no. 10, pp. 1219 - 1225.

Evans, S. J. & Sayers, A. R. (2000): A longitudinal study of campylobacter infection of broiler flocks in Great Britain. *Preventive Veterinary Medicine 46*, no. 3, pp. 209 - 223.

Eyles, R. (1999): Campylobacteriosis in Aoteoroa New Zealand: the usefulness of GIS to a current public health problem. Retrieved: 14/07/2006, from: http://www.business.otago.ac.nz/SIRC05/conferences/1999/32_Eyles.pdf.

Eyles, R.; Buck, O.; Niyogi, D. & Townsend, C. (2002): Land use, water quality and human health risk: results from two small agricultural catchments. Retrieved: 19/05/2007, from: http://www.business.otago.ac.nz/SIRC05/conferences/2002_SIRC/05_Eyles.pdf.

Eyles, R.; Niyogi, D.; Townsend, C.; Benwell, G. & Weinstein, P. (2003): Spatial and temporal patterns of Campylobacter contamination underlying public health risk in the Taieri River, New Zealand. *Journal of Environmental Quality 32*, no. 5, pp. 1820 - 1828.

EZDRSC (2007): Enteric Zoonotic Disease Research Steering Committee. Retrieved: 24/02/2007, from: http://www.zoonosesresearch.org.nz/.

Friedman, C. R.; Hoekstra, R. M.; Samuel, M.; Marcus, R.; Bender, J.; Shiferaw, B.; Reddy, S.; Ahuja, S. D.; Helfrick, D. L.; Hardnett, F.; Carter, M.; Anderson, B. & Tauxe, R. V. (2004): Risk factors for sporadic Campylobacter infection in the United States: A case-control study in FoodNet sites. *Clinical Infectious Diseases 38*, no. 3, pp. 285 - 296.

Frost, J. A.; Gillespie, I. A. & O'Brien, S. J. (2002): Public health implications of campylobacter outbreaks in England and Wales, 1995-9: epidemiological and microbiological investigations. *Epidemiology and Infection 128*, no. 2, pp. 111 - 118.

Fülgraff, G. (2001): Schnittstellen der Umweltmedizin. *Umweltmedizin in Forschung und Praxis 6*, no. 6, pp. 303 - 308.

Gallay, A.; De Valk, H.; Cournot, M.; Ladeuil, B.; Hemery, C.; Castor, C.; Bon, F.; Megraud, F.; Le Cann, P. & Desenclos, J. C. (2006): A large multi-pathogen waterborne community outbreak linked to faecal contamination of a groundwater system, France, 2000. *Clinical Microbiology and Infection 12*, no. 6, pp. 561 - 570.

Garg, A. X.; Marshall, J.; Salvadori, M.; Thiessen-Philbrook, H. R.; Macnab, J.; Suri, R. S.; Haynes, R. B.; Pope, J.; Clark, W.; Collins, S.; Howard, J.; Mahon, J.; Matsell, D.; Moist, L.; Ray, J. & Rosas-Arellano, P. (2006): A gradient of acute gastroenteritis was characterized, to assess risk of long-term health sequelae after drinking bacterial-contaminated water. *Journal of Clinical Epidemiology 59*, no. 4, pp. 421 - 428.

Gatrell, A. C. (2002): *Geographies of Health. An Introduction.* Blackwell, Malden, Oxford, Victoria.

Gillespie, I. A.; O'Brien, S. J.; Adak, G. K.; Tam, C. C.; Frost, J. A.; Bolton, F. J. & Tompkins, D. S. (2003): Point source outbreaks of Campylobacter jejuni infection--are they more common than we think and what might cause them? *Epidemiology and Infection 130*, no. 3, pp. 367 - 375.

Gilpin, B.; Cornelius, A.; Robson, B.; Boxall, N.; Ferguson, A.; Nicol, C. & Henderson, T. (2006): Application of pulsed-field gel electrophoresis to identify potential outbreaks of campylobacteriosis in New Zealand. *J Clin Microbiol 44*, no. 2, pp. 406 - 412.

Giskes, K.; Turrell, G.; Patterson, C. & Newman, B. (2002): Socio-economic differences in fruit and vegetable consumption among Australian adolescents and adults. *Public Health Nutrition 5*, no. 5, pp. 663 - 669.

Gochfeld, M. (2005): Chronologic history of occupational medicine. *Journal of Occupational and Environmental Medicine 47*, no. 2, pp. 96 - 114.

Gordis, L. (2000): *Epidemiology.* W.B. Saunders, Philadelphia.

Gottlieb, N. H.; Loukas, A.; Corrao, M.; McAlister, A.; Snell, C. & Huang, P. P. (2004): Minors' tobacco possession law violations and intentions to smoke: implications for tobacco control. *Tobacco Control 13*, no. 3, pp. 237 - 243.

Greenland, S. (1998): Introduction to Regression Modeling. In: Rothman, K. J. & Greenland, S. (Eds.): *Modern Epidemiology.* Lippincott Williams & Wilkins, Philadelphia, Baltimore, New York, pp. 401 - 432.

Hales, S.; Black, W.; Skelly, C.; Salmond, C. & Weinstein, P. (2003): Social deprivation and the public health risks of community drinking water supplies in New Zealand. *Journal of Epidemiology and Community Health 57*, no. 8, pp. 581 - 583.

Hänninen, M. L.; Haajanen, H.; Pummi, T.; Wermundsen, K.; Katila, M. L.; Sarkkinen, H.; Miettinen, I. & Rautelin, H. (2003): Detection and typing of Campylobacter jejuni and Campylobacter coli and analysis of indicator organisms in three waterborne outbreaks in Finland. *Applied and Environmental Microbiology 69*, no. 3, pp. 1391 - 1396.

Hatch, M. C.; Wallenstein, S.; Beyea, J.; Nieves, J. W. & Susser, M. (1991): Cancer rates after the Three Mile Island nuclear accident and proximity of residence to the plant. *American Journal of Public Health 81*, no. 6, pp. 719 - 724.

Hearnden, M.; Skelly, C.; Eyles, R. & Weinstein, P. (2003): The regionality of campylobacteriosis seasonality in New Zealand. *International Journal of Environmental Health Research 13*, no. 4, pp. 337 - 348.

Hearnden, M.; Weinstein, P.; Firns, P.; Skelly, C.; Salmond, C. & Benwell, G. (1998): An ecological approach to environmental sources of campylobacteriosis in New Zealand. Retrieved: 28/07/2006, from: http://www.business.otago.ac.nz/SIRC05/conferences/1998/16_Hearnden.pdf.

Hertz-Picciotto, I. (1998): Environmental Epidemiology. In: Rothman, K. J. & Greenland, S. (Eds.): *Modern Epidemiology.* Lippincott Williams & Williams, Philadelphia, Baltimore, New York, pp. 555 - 583.

Hood, A. M.; Pearson, A. D. & Shahamat, M. (1988): The extent of surface contamination of retailed chickens with Campylobacter jejuni serogroups. *Epidemiology and Infection 100*, no. 1, pp. 17 - 25.

Hoque, M. E.; Hope, V. T.; Scragg, R. & Graham, J. (2005): Under-notification of giardiasis in Auckland, New Zealand: a capture-recapture estimation. *Epidemiology and Infection 133*, no. 1, pp. 71 - 79.

Hrudey, S. E. & Hrudey, E. J. (2004): *Safe Drinking Water. Lessons from Recent Outbreaks in Affluent Nations.* IWA Publishing, London.

Hudson, J. A.; Nicol, C.; Wright, J.; Whyte, R. & Hasell, S. K. (1999): Seasonal variation of Campylobacter types from human cases, veterinary cases, raw chicken, milk and water. *Journal of Applied Microbiology 87*, no. 1, pp. 115 - 124.

Ikram, R.; Chambers, S.; Mitchell, P.; Brieseman, M. A. & Ikam, O. H. (1994): A case control study to determine risk factors for campylobacter infection in Christchurch in the summer of 1992-3. *New Zealand Medical Journal 107*, no. 988, pp. 430 - 432.

Johnston, R. J. (1985): Places matter. *Irish Geography 18*, no. 2, pp. 58 - 63.

Jones, K. (2001): Campylobacters in water, sewage and the environment. *Symposium Series Societey for Applied Microbiology*, no. 30, pp. 68S-79S.

Jusatz, H. (1987): Grundlagen und Grundbegriffe der Geomedizin. *Erdkundliches Wissen 70*, no. n.s., pp. 11 - 24.

Kalipeni, E.; Craddock, S. & Oppong, J. (2004): *HIV and AIDS in Africa beyond epidemiology.* Blackwell Publishing, Malden, Oxford, Carlton.

Kapperud, G.; Espeland, G.; Wahl, E.; Walde, A.; Herikstad, H.; Gustavsen, S.; Tveit, I.; Natas, O.; Bevanger, L. & Digranes, A. (2003): Factors associated with increased and decreased risk of Campylobacter infection: a prospective case-control study in Norway. *American Journal of Epidemiology 158*, no. 3, pp. 234 - 242.

Kearns, R. A. & Joseph, A. E. (1993): Space in Its Place - Developing the Link in Medical Geography. *Social Science and Medicine 37*, no. 6, pp. 711 - 717.

Keener, K. M.; Bashor, M. P.; Curtis, P. A.; Sheldon, B. W. & Kathariou, S. (2004): Comprehensive Review of *Campylobacter* and Poultry Processing. *Comprehensive Reviews in Food Science and Food Safety 3*, no. n/a, pp. 105 - 116.

Khan, R.; Voice, T. & Phillips, D. (2005): Developing National Environmental Health Indicators for New Zealand. Client Report FW 0570. Retrieved 12/07/2006, from: http://www.surv.esr.cri.nz/PDF_surveillance/EHI/EHI_AnnualReport_2005.pdf.

Kingham, S. P.; Gatrell, A. C. & Rowlingson, B. (1995): Testing for clustering of health events within a geographical information system framework. *Environment and Planning A 27*, no. 5, pp. 809 - 821.

Kistemann, T. (1997): Waterborne infections - Risks in highly developed water supply structures. *Geographische Rundschau 49*, no. 4, pp. 210 - 215.

Kistemann, T.; Claßen, T.; Koch, C.; Dangendorf, F.; Fischeder, R.; Gebel, J.; Vacata, V. & Exner, M. (2002): Microbial load of drinking water reservoir tributaries during extreme rainfall and runoff. *Applied and Environmental Microbiology 68*, no. 5, pp. 2188 - 2197.

Kistemann, T.; Leisch, H. & Schweikart, J. (1997): Geomedizin und Medizinische Geographie. Entwicklung und Perspektiven einer "old partnership". *Geographische Rundschau 49*, no. 4, pp. 198 - 203.

Kistemann, T. & Queste, A. (2004): GIS and Communicable Disease Control. In: Maheswaran, R. & Craglia, M. (Eds.): *GIS in public health practice.* CRC Press, Boca Raton, FL, pp. 71 - 90.

Kleinschmidt, I.; Sharp, B.; Mueller, I. & Vounatsou, P. (2002): Rise in malaria incidence rates in South Africa: a small-area spatial analysis of variation in time trends. *American Journal of Epidemiology 155*, no. 3, pp. 257 - 264.

Koch, T. & Denike, K. (2001): GIS approaches to the problem of disease clusters: A brief commentary. *Social Science and Medicine 52*, no. 11, pp. 1751 - 1754.

Kovats, R. S.; Edwards, S. J.; Charron, D.; Cowden, J.; D'Souza, R. M.; Ebi, K. L.; Gauci, C.; Gerner-Smidt, P.; Hajat, S.; Hales, S.; Pezzi, G. H.; Kriz, B.; Kutsar, K.; McKeown, P.; Mellou, K.; Menne, B.; O'Brien, S.; van Pelt, W. & Schmid, H. (2005): Climate variability and campylobacter infection: an international study. *International Journal of Biometeorology 49*, no. 4, pp. 207 - 214.

Kremer, A. (2004): Urbane Umwelt und Gesundheit. Exposition und Risikowahrnehmung vulnerabler Bevölkerungsgruppen in Pondycherry, Indien. Mathematisch-Naturwissenschaftliche Fakultät der Universität Bonn. Retrieved: 24/6/2006, from: http://hss.ulb.uni-bonn.de/diss_online/math_nat_fak/2004/kremer_anne/index.htm.

Kulldorff, M.; Heffernan, R.; Hartman, J.; Assuncao, R. & Mostashari, F. (2005): A space-time permutation scan statistic for disease outbreak detection. *PLoS medicine. 2*, no. 3, pp. 216 - 224.

Kulldorff, M.; Song, C.; Gregorio, D.; Samociuk, H. & DeChello, L. (2006): Cancer map patterns: Are they random or not? *American Journal of Preventive Medicine 30*, no. 2 SUPPL., pp. S37 - S49.

Kurland, L. T. & Reed, D. (1988): Geographic and climatic aspects of multiple sclerosis: a review of current hypotheses. In: Buck, C.; Llopis, A.; Nájera, E. & Terris, M. (Eds.): *The challenge of epidemiology. Issues and selected readings.* WHO printing, Washington, D.C., pp. 393 - 401.

Lake, H.; Hudson, A.; Cressey, P. & Nortje, G. (2003): Risk Profile: Campylobacter Jejuni/ Coli in Poultry (Whole and Pieces). Client Report FW0109. Retrieved 24/07/2006, from: http://www.nzfsa.govt.nz/science/risk-profiles/campylobacter.pdf.

Lawson, A. & Williams, F. L. R. (2001): *An introductory guide to disease mapping.* John Wiley, Chichester, New York.

Leathwick, J. R.; Wilson, G. & Stephens, R. T. T. (2002): Climate Surfaces for New Zealand. Landcare Research Contract Report: LC9798/126. Retrieved: 19/02/2007, from: http://www.landcareresearch.co.nz/databases/lenz/downloads/Climate_Surfaces_for_New_Zealand_revised.pdf.

Leyland, A. H. & Groenewegen, P. P. (2003): Multilevel modelling and public health policy. *Scandinavian Journal of Public Health 31*, no. 4, pp. 267 - 274.

Lindstrom, M.; Hanson, B. S.; Wirfalt, E. & Ostergren, P. O. (2001): Socioeconomic differences in the consumption of vegetables, fruit and fruit juices. The influence of psychosocial factors. *European Journal of Public Health 11*, no. 1, pp. 51 - 59.

Louis, V. R.; Gillespie, I. A.; O'Brien, S. J.; Russek-Cohen, E.; Pearson, A. D. & Colwell, R. R. (2005): Temperature-driven campylobacter seasonality in England and Wales. *Applied and Environmental Microbiology 71*, no. 1, pp. 85 - 92.

Luke, D.; Esmundo, E. & Bloom, Y. (2000): Smoke signs: patterns of tobacco billboard advertising in a metropolitan region. *Tobacco Control 9*, no. 1, pp. 16 - 23.

Macfie, B. P. & Nufrio, P. M. (2006): *Applied statistics for public policy.* M.E. Sharpe Inc., Armonk, N.Y.

MAF (2002a): Livestock by Territorial Authority. Retrieved: 05/08/2006, from: http://www.maf.govt.nz/statistics/primaryindustries/livestock/summary-tables/index.htm.

MAF (2002b): Poultry Statistic. Retrieved: 03/08/2006, from: http://www.maf.govt.nz/statistics/primaryindustries/livestock/poultry/index.htm.

MAF (2003): Pig and poultry industries. Retrieved: 05/05/2007, from: http://www.maf.govt.nz/mafnet/rural-nz/profitability-and-economics/contribution-of-land-based-industries-nz-economic-growth/contribution17.htm.

MAF (2006): A History of the Stock Unit System. Retrieved: 14/07/2006, from: http://www.maf.govt.nz/mafnet/rural-nz/sustainable-resource-use/best-management-practices/reassessment-of-the-stock-management-system/re-assessment-of-stock-unit-system03.htm.

Manicas, P. T. (2006): *A realist philosophy of social science: explanation and understanding.* Cambridge University Press, Cambridge, UK ; New York.

Martin, R. M. (2005): Epidemiological study designs for health care research and evaluation. In: Bowling, A. & Ebrahim, S. (Eds.): *Handbook of health research methods. Investigation, measurement and analysis.* Open University Press, Maidenhead, Berkshire, New York, pp. 98 - 163.

May, J. M. (1951): The geography of pathology. *Scientific Monthly 72*, no. 2, pp. 128 - 129.

Mayer, J. D. (1981): Geographical Clues About Multiple-Sclerosis. *Annals of the Association of American Geographers 71*, no. 1, pp. 28 - 39.

Mayer, J. D. (1982): Relations between two traditions of medical geography: health systems planning and geographical epidemiology. *Progress in Human Geography 6*, no. 2, pp. 216 - 230.

Mayer, J. D. (1990): The centrality of medical geography to human geography: the traditions of geographical and medical geographical thought. *Norsk Geografisk Tidsskrift 44*, no. 4, pp. 175 - 187.

Mayer, J. D. (1996): The political ecology of disease as one new focus for medical geography. *Progress in Human Geography 20*, no. 4, pp. 441 - 456.

McBride, G. B.; Till, D. G.; Ryan, T.; Ball, A.; Lewis, G.; Palmer, S. & Weinstein, P. (2002): Freshwater Microbiology Research Programme. Pathogen Occurence and Human Health Risk. Assessment Analysis. Retrieved: 19/05/2007, from: http://www.mfe.govt.nz/publications/water/freshwater-microbiology-nov02/freshwater-microbiology-main-nov02.pdf.

McNicholas, A. M.; Bates, M.; Kiddle, E. & Wright, J. (1995): Is New Zealand's recent increase in campylobacteriosis due to changes in laboratory procedures? A survey of 69 medical laboratories. *New Zealand Medical Journal 108*, no. 1011, pp. 459 - 461.

Meade, M. & Earickson, R. J. (2002): *Medical Geography.* Guilford Press, New York, London.

Michaud, S.; Ménard, S. & Arbeit, R. D. (2004): Campylobacteriosis, Eastern Townships, Québec. *Emerging Infectious Diseases 10*, no. 10, pp. 1844 - 1847.

Miles, S.; Braxton, D. S. & Frewer, L. J. (1999): Public perceptions about microbiological hazards in food. *British Food Journal 101*, no. 10, pp. 744 - 762.

Mitchell, P. (2000): Using GIS to determine if there is an Association between Water Quality and Enteric Disease in Canterbury. Retrieved: 28/07/2006, from: http://www.business.otago.ac.nz/SIRC05/conferences/2000/05_mitchell.pdf.

Mitchell, P.; Ambrose, C. & Williams, D. (2002): Disease Report Campylobacteriosis. CROWN PUBLIC HEALTH Number 6. Retrieved: 09/08/2006, from: http://www.cpublichealth.co.nz/Files/DiseaseReport6.pdf.

MoH (2004): *Annual Review of the Microbiological and Chemical Quality of Drinking-Water in New Zealand 2004.* Ministry of Health, Wellington.

MoH (2005): Atlas of Cancer Mortality in New Zealand 1994 - 2000. Retrieved: 03/10/2006, from: http://www.moh.govt.nz/moh.nsf/238fd5fb4fd051844c256669006aed57/cdbc682ae572e9a2cc256fda0072860f?OpenDocument.

MoH (2005 -a): *Draft Guidelines for Drinking-water Quality Management for New Zealand 2005. Second edition.* Ministry of Health, Wellington.

MoH (2005 -b): *Drinking-water standards for New Zealand 2005.* Ministry of Health, Wellington.

MoH (2005 -c): *Register of Community Drinking-Water Supplies in New Zealand.* Ministry of Health, Wellington.

MoH (2006): *Register of Community Drinking-Water Supplies in New Zealand.* Ministry of Health, Wellington.

Moon, G. & Gould, M. (2000): *Epidemiology: An Introduction.* Open University Press, Buckimgham, Philadelphia.

Morgan, C. (2002): Interpreting spatial and temporal patterns in South Australian Campylobacter Surveillance Data: 1990 - 1999. Retrieved: 09/08/2006, from: http://www.business.otago.ac.nz/SIRC05/conferences/2002_GeoHealth/11_morgan.pdf.

Morgenstern, H. (1995): Ecologic studies in epidemiology: Concepts, principles, and methods. *Annual Review of Public Health 16*, no. 3, pp. 61 - 81.

Morgenstern, H. & Thomas, D. (1993): Principles of study design in environmental epidemiology. *Environmental Health Perspectives 101*, no. SUPPL. 4, pp. 23 - 38.

Morris, S. E. & Wakefield, J. C. (2000): Assessment of disease risk in relation to a pre-specified source. In: Elliot, P.; Wakefield, J. C.; Best, N. G. & Briggs, D. J. (Eds.): *Spatial Epidemiology. Methods and Applications.* Oxford University Press, Oxford, New York, pp. 153 - 184.

Morrison, D. & Smith, D. (2004): *Campylobacteriosis in Christchurch City.* Community and Public Health, Christchurch.

National Health Committee (2002): The Health of People and Communities. The effect of environmental factors on the health of New Zealanders. Report to the Minister of Health from the Public Health Advisory Committee. Retrieved 23/08/2006, from: http://www.nhc.govt.nz/phac/publications/Health%20of%20People.pdf.

National Institute of Justice (2004): *CrimeStat® III. A Spatial Statistics Program for the Analysis of Crime Incident Locations.* National Institute of Justice, Washington.

Nelson, W. & Harris, B. (2006): Flies, fingers, fomites, and food. Campylobacteriosis in New Zealand--food-associated rather than food-borne. *New Zealand Medical Journal 119*, no. 1240, pp. U2128.

New Zealand Government (2007): $2.4 million to fund 10 new research projects. Retrieved: 09/05/2007, from: http://www.scoop.co.nz/stories/PA0702/S00537.htm.

Newell, D. G.; Carnevale, R.; Thornsberry, C. & Gonder, E. (2002): The ecology of Campylobacter jejuni in avian and human hosts and in the environment. *International Journal of Infectious Diseases 6*, no. SUPPL. 3, pp. 3S16 - 3S21.

Nichols, G. L. (2005): Fly transmission of Campylobacter. *Emerging Infectious Diseases 11*, no. 3, pp. 361 - 364.

Nielsen, E. M.; Fussing, V.; Engberg, J.; Nielsen, N. L. & Neimann, J. (2006): Most Campylobacter subtypes from sporadic infections can be found in retail poultry products and food animals. *Epidemiology and Infection 134*, no. 4, pp. 758 - 767.

Norström, M.; Hofshagen, M.; Stavnes, T.; Schau, J.; Lassen, J. & Kruse, H. (2006): Antimicrobial resistance in Campylobacter jejuni from humans and broilers in Norway. *Epidemiology and Infection 134*, no. 1, pp. 127 - 130.

Norušis, M. J. (2004): *SPSS 12.0. Guide to Data Analysis.* Prentice Hall, Upper Saddle River, New Jersey.

Nygård, K.; Andersson, Y.; Røttingen, J. A.; Svensson, Å.; Lindbäck, J.; Kistemann, T. & Giesecke, J. (2004): Association between environmental risk factors and campylobacter infections in Sweden. *Epidemiology and Infection 132*, no. 2, pp. 317 - 325.

Nylen, G.; Dunstan, F.; Palmer, S. R.; Andersson, Y.; Bager, F.; Cowden, J.; Feierl, G.; Galloway, Y.; Kapperud, G.; Megraud, F.; Molbak, K.; Petersen, L. R. & Ruutu, P. (2002): The seasonal distribution of campylobacter infection in nine European countries and New Zealand. *Epidemiology and Infection 128*, no. 3, pp. 383 - 390.

NZFSA (2005): New Zealand Food Safety Authority. A quantitative study. Retrieved: 27/03/2007, from: http://www.nzfsa.govt.nz/science/research-projects/umr-research/nzfsa-quantitativestudy-may05-2.pdf.

NZFSA (2006): *Campylobacter* in poultry - Risk Management Strategie 2006 - 2009. Retrieved: 09/05/2007, from: http://www.nzfsa.govt.nz/publications/news-current-issues/nzfsa-campylobacter-strategy-nov-2006.pdf.

NZFSA (2007): Grant will help fund *Campylobacter* research. Retrieved: 09/05/2005, from: http://www.nzfsa.govt.nz/publications/media-releases/2007/20070411.htm#P9_108.

NZPHO (2006): Notifiable Disease and Environmental Health Indicator Dataset. Retrieved: 07/072006, from: http://www.nzpho.org.nz/data_available.asp.

Osborn, J. W. (2000): Prediciton in multiple regression. *Practical Assessment, Research & Evaluation 7*, no. 2, pp. 15 - 17.

Pallant, J. (2005): *SPSS survival manual : a step by step guide to data analysis using SPSS for Windows (Version 12).* Allen & Unwin, Crows Nest, N.S.W.

Pearce, J. & Dorling, D. (2006): Increasing geographical inequalities in health in New Zealand, 1980-2001. *International Journal of Epidemiology 35*, no. 3, pp. 597 - 603.

Pearce, J.; Witten, K. & Bartie, P. (2006): Neighbourhoods and health: A GIS approach to measuring community resource accessibility. *Journal of Epidemiology and Community Health 60*, no. 5, pp. 389 - 395.

Pedhazur, E. J. (1997): *Multiple Regression in Behavioral Research. Explanation and Prediction.* Harcourt Brace College Publishers, Fort Worth, Philadelphia, San Diego.

Pedhazur, E. J. & Schmelkin, L. P. (1991): *Measurement, design, and analysis : an integrated approach.* Lawrence Erlbaum Associates, Hillsdale, N.J.

Percival, S.; Chalmers, R.; Embrey, M.; Hunter, P.; Sellwood, J. & Wyn-Jones, P. (2004): *Microbiology of Waterborne Diseases.* Academic Press, Amsterdam, Boston, Heidelberg.

Peterson, M. C. (1994): Clinical aspects of Campylobacter jejuni infections in adults. *Western Journal of Medicine 161*, no. 2, pp. 148 - 152.

Phillips, C. V. & Goodman, K. J. (2004): The missed lessons of Sir Austin Bradford Hill. *Epidemiologic Perspectives & Innovation 1*, no. 1, pp. 3 - 7.

Physicians Committee for Responsible Medicine (2001): Americans Don't Realize Poultry and Meat Are Often Contaminated with Feces. Retrieved: 07/05/2007, from: http://www.pcrm.org/news/health010829survey.html.

Pond, K. (2005): *Water Recreation and Disease. Plausibility of Associated Infections: Acute Effects, Sequelae and Mortality.* IWA Publishing, London.

Potter, R. C.; Kaneene, J. B. & Hall, W. N. (2003): Risks Factors for Sporadic Campylobacter jejuni Infections in Rural Michigan: A Prospective Case-Control Study. *American Journal of Public Health 93*, no. 12, pp. 2118 - 2123.

Public Health Advisory Committee (2004): Emerging Issues for Public Health in New Zealand. Retrieved: 14/07/2006, from: http://www.phac-aspc.gc.ca/php-psp/pdf/emerging_issues_for_public_health_in_new%20_zaland.pdf.

Public Health Surveillance (2007): Information for New Zealand Public Health Action. Retrieved: 24/02/2007, from: http://www.surv.esr.cri.nz/index.php.

Reiersen, J.; Briem, H.; Hardardottir, H.; Gunnarsson, E.; Georgsson, F.; Gudmundsdottir, E. & Kristinsson, K. (2002): Human Campylobacteriosis Epidemic in Iceland 1998 - 2000 and Effect of Interventions aimed at Poultry and Humans. Retrieved: 07/08/2006, from: http://www.fao.org/DOCREP/MEETING/004/AB520E.HTM.

Rezaeian, M.; Dunn, G.; St Leger, S. & Appleby, L. (2007): Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. *Journal of Epidemiology and Community Health 61*, no. 2, pp. 98 - 102.

Rican, S.; Salem, G. & Jougla, E. (1999): Geographic differences of bronchopulmonary cancer mortality in France and spatial scales of analysis: significance of scale change in health geography. *Revue d'epimiologie et de sante publique 47*, no. 5, pp. 411 - 420.

Robert Koch-Institute (2005): *Campylobacter* Infections. Retrieved: 16/08/2006, from: http://www.rki.de/cln_049/nn_196658/DE/Content/Infekt/EpidBull/Merkblaetter/Ratgeber__Mbl__Campylobacter.html.

Rosef, O.; Rettedal, G. & Lageide, L. (2001): Thermophilic campylobacters in surface water: a potential risk of campylobacteriosis. *International Journal of Environmental Health Research 11*, no. 4, pp. 321 - 327.

Rothman, K. J. (1986): *Modern epidemiology.* Little Brown, Boston.

Rothman, K. J. & Greenland, S. (1998): The Emergence of Modern Epidemiology. In: Rothman, K. J. & Greenland, S. (Eds.): *Modern Epidemiology.* 2 ed. Lippincott Williams & Wilkins, Philadelphia, pp. 3 - 6.

Rothman, K. J. & Greenland, S. (2005): Causation and Causal Inference in Epidemiology. *American Journal of Epidemiology 95*, no. S1, pp. S144 - S150.

Sabel, C. E.; Boyle, P. J.; Löytönen, M.; Gatrell, A. C.; Jokelainen, M.; Flowerdew, R. & Maasilta, P. (2003): Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. *American Journal of Epidemiology 157*, no. 10, pp. 898 - 905.

Sabel, C. E.; Gatrell, A. C.; Löytönen, M.; Maasilta, P. & Jokelainen, M. (2000): Modelling exposure opportunities: Estimating relative risk for motor neurone disease in Finland. *Social Science and Medicine 50*, no. 7-8, pp. 1121 - 1137.

Sabel, C. E. & Löytönen, M. (2004): Clustering of Disease. In: Maheswaran, R. & Craglia, M. (Eds.): *GIS in Public Health Practice.* CRC Press, Boca Raton, FL, pp. 51 - 70.

Said, B.; Wright, F.; Nichols, G. L.; Reacher, M. & Rutter, M. (2003): Outbreaks of infectious diseases associated with private drinking water supplies in England and Wales 1970 - 2000 *Epidemiology and Infection 130*, no. 3, pp. 497 - 479.

Salmond, C. & Crampton, P. (2002): NZDep2001 Index of Deprivation. User's Manual. Retrieved: 14/07/2006, from: http://www.moh.govt.nz/moh.nsf/Files/phi-users-manual/$file/phi-users-manual.pdf.

Salmond, C.; Crampton, P.; King, P. & Waldegrave, C. (2006): NZiDep: A New Zealand index of socioeconomic deprivation for individuals. *Social Science and Medicine 62*, no. 6, pp. 1474 - 1485.

Salmond, C. & Salmond, P. (2002): NZDep2001 Index of Deprivation. Retrieved: 13/07/2006, from: http://www.moh.govt.nz/moh.nsf/Files/phi-research-report/$file/phi-research-report.pdf.

Sandberg, M. M.; Nygård, K. K.; Meldal, H. H.; Valle, P. S.; Kruse, H. H. & Skjerve, E. E. (2006): Incidence trend and risk factors for Campylobacter infections in humans in Norway. *BMC Public Health 6*, no. 1, pp. 179 - 205.

Savill, M.; Hudson, A.; Devane, M.; Garrett, N.; Gilpin, B. & Ball, A. (2003): Elucidation of potential transmission routes of Campylobacter in New Zealand. *Water Science and Technology 47*, no. 3, pp. 33 - 38.

Savill, M. G.; Hudson, J. A.; Ball, A.; Klena, J. D.; Scholes, P.; Whyte, R. J.; McCormick, R. E. & Jankovic, D. (2001): Enumeration of Campylobacter in New Zealand recreational and drinking waters. *Journal of Applied Microbiology 91*, no. 1, pp. 38 - 46.

Schærström, A. (1996): *Pathogenic paths? A time geographical approach in medical geography.* Lund University Press, Lund.

Schildt, M.; Savolainen, S. & Hanninen, M. L. (2006): Long-lasting Campylobacter jejuni contamination of milk associated with gastrointestinal illness in a farming family. *Epidemiology and Infection 134*, no. 2, pp. 401 - 405.

Schonwalder, C. & Olden, K. (2003): Environmental health moves into the 21st century. *International Journal of Hygiene and Environmental Health 206*, no. 4 - 5, pp. 263 - 267.

Scott, W. G.; Scott, H. M.; Lake, R. J. & Baker, M. G. (2000): Economic cost to New Zealand of foodborne infectious disease. *New Zealand Medical Journal 113*, no. 1113, pp. 281 - 284.

Shannon, G. W.; Pyle, G. F. & Bashshur, R. (1991): *The geography of AIDS : origins and course of an epidemic.* Guilford Press, New York.

Shaw, G. & Wheeler, D. (1985): *Statistical techniques in geographical analysis.* Wiley, Chichester; New York.

Shohaimi, S.; Welch, A.; Bingham, S.; Luben, R.; Day, N.; Wareham, N. & Khaw, K. T. (2004): Residential area deprivation predicts fruit and vegetable consumption independently of individual educational level and occupational social class: a cross sectional population study in the Norfolk cohort of the European Prospective Investigation into Cancer (EPIC-Norfolk). *Journal of Epidemiology and Community Health 58*, no. 8, pp. 686 - 691.

Skelly, C.; Black, W.; Hearnden, M.; Eyles, R. & Weinstein, P. (2002): Disease surveillance in rural communities is compromised by address geocoding uncertainty: a case study of campylobacteriosis. *Australian Journal of Rural Health 10*, no. 2, pp. 87 - 93.

Sneyd, E. & Baker, M. (2003): Infectious diseases in New Zealand: 2002 annual surveillance summary. Retrieved: 13/02/2007, from: http://www.surv.esr.cri.nz/PDF_surveillance/AnnSurvRpt/2002AnnualSurvRpt.pdf.

SNZ (2001): *Census of Population and Dwellings Statistics. Territorial Authority Areas of Usual Residence by Occupation. Data received via email, Christchurch, New Zealand. Date: 24/08/2006.*

SNZ (2004): Territorial Authority. Retrieved: 24/02/2004, from: http://www.stats.govt.nz/statistical-methods/classifications/Territorial-authority-2004a.htm.

SNZ (2006): New Zealand: An Urban/ Rural Profile. Retrieved: 20/02/2007, from: http://www.stats.govt.nz/urban-rural-profiles/default.htm.

Sopwith, W.; Ashton, M.; Frost, J. A.; Tocque, K.; O'Brien, S.; Regan, M. & Syed, Q. (2003): Enhanced surveillance of campylobacter infection in the North West of England 1997-1999. *Journal of Infection 46*, no. 1, pp. 35 - 45.

Sopwith, W.; Birtles, A.; Matthews, M.; Fox, A.; Gee, S.; Painter, M.; Regan, M.; Syed, Q. & Bolton, E. (2006): Campylobacter jejuni multilocus sequence types in humans, northwest England, 2003-2004. *Emerging Infectious Diseases 12*, no. 10, pp. 1500 - 1507.

Stanley, K. & Jones, K. (2003): Cattle and sheep farms as reservoirs of Campylobacter. *Journal of Applied Microbiology 94*, no. 3, pp. 104 - 113.

Stefanogiannis, N.; Lawes, C. M.; Turley, M.; Tobias, M.; Hoorn, S. V.; Mhurchu, C. N. & Rodgers, A. (2005): Nutrition and the burden of disease in New Zealand: 1997-2011. *Public Health Nutrition 8*, no. 4, pp. 395 - 401.

Stehrgreen, J. K.; Nicholls, C.; McEwan, S.; Payne, A. & Mitchell, P. (1991): Waterborne Outbreak of Campylobacter-Jejuni in Christchurch - the Importance of a Combined Epidemiologic and Microbiologic Investigation. *New Zealand Medical Journal 104*, no. 918, pp. 356 - 358.

Tauxe, R. V.; Hargrett-Bean, N.; Patton, C. M. & Wachsmuth, K. (1988): Campylobacter Isolates in the United States, 1982 - 1986. Retrieved: 27/10/2006, from: http://www.cdc.gov/mmwr/preview/mmwrhtml/00001764.htm.

TeckLok, W.; On, S. L. W. & Michie, H. (2006): Campylobacter in New Zealand: Reservoirs, Sources and the Labyrinth of Transmission Routes. *New Zealand Journal of Environmental Health 29*, no. 2, pp. 1 - 6.

Thagard, P. (1997): The concept of disease: Structure and Change. Retrieved: 21/08/2006, from: http://cogsci.uwaterloo.ca/Articles/Pages/Concept.html.

The Lifestyleblock.co.nz (2006): Livestock Units or Ewe Equivalents. Retrieved: 14/07/2006, from: http://www.lifestyleblock.co.nz/articles/general/14_livestock_units.htm.

Thompson, S.; Grüner, I. & Gapare, N. (2003): *New Zealand Land Cover Database Version 2. Illustrated Guide to Target Classes.* Ministry for the Environment, Wellington.

Thornley, C.; McDowell, R.; Lopez, L. & Baker, M. (2002): Annual summary of outbreaks in New Zealand 2001. Client Report FW 0218. Retrieved 23/07/2006, from: http://www.surv.esr.cri.nz/PDF_surveillance/AnnSurvRpt/2001OutbreakRpt.pdf.

Till, D. G. & McBride, G. B. (2005): Potential public health risk of *Campylobacter* and other zoonotic waterborne infections in New Zealand. In: Cotruvo, J., Dufour,A., Rees,G., Bartram,J., Carr,R., Cliver,D.O., Craun,G.F. (Eds.): *Waterborne Zoonoses: Identification, Causes and Control.* IWA Publishing, Essex, UK, pp. 191 - 207.

Tobias, M.; Jackson, G.; Yeh, L. C. & Huang, K. (2007): Do healthy and unhealthy behaviours cluster in New Zealand? *Austalian and New Zealand Journal of Public Health 31*, no. 2, pp. 155 - 163.

Tobias, M. I. & Cheung, J. (2003): Monitoring health inequalities: life expectancy and small area deprivation in New Zealand. *Population Health Metrics 1*, no. 1, pp. 1 - 11.

Todar, K. (2003): Todar's Online Textbook of Bacteriology. Retrieved: 24/10/2006, from: http://textbookofbacteriology.net/Listeria.html.

Tunstall, H. V.; Shaw, M. & Dorling, D. (2004): Places and health. *Journal of Epidemiology and Community Health 58*, no. 1, pp. 6 - 10.

U.S. Department of Health and Human Services (1998): An Ensemble of Definitions of Environmental Health. Retrieved: 14/07/2006, from: http://www.health.gov/environment/definitionsofenvhealth/ehdef2.htm.

van den Berg, N. & von der Ahé, K.-R. (1997): Geoinformationssysteme in der Epidemiologie. *Karthographische Nachrichten 47*, no. 2, pp. 52 - 59.

van der Mei, I. A. F.; Ponsonby, A.-L.; Dwyer, T.; Blizzard L; Simmons, R.; Taylor, B. V.; Butzkueven, H. & Kilpatrick, T. (2003): Past exposure to sun, skin phenotype, and risk of multiple sclerosis: case-control study. *BMJ 327*, no. 1468-5833 (electronic), pp. 316 - 320.

Wakefield, J. C.; Best, N. G. & Waller, L. (2000): Clustering, cluster detection, and spatial variation in risk. In: Elliott, P.; Wakefield, J. C.; Best, N. G. & Briggs, D. J. (Eds.): *Spatial Epidemiology. Methods and Applications.* Oxford University Press, Oxford, New York, pp. 128 - 152.

Walford, N. (1995): *Geographical data analysis.* J. Wiley & Sons, Chichester, New York.

Wall, P. A. & Devine, O. J. (2000): Interactive analysis of the spatial distribution of disease using a geographic information system. *Journal of Geographical Systems 2*, no. 3, pp. 243 - 256.

Weinstein, P.; Russell, N. & Woodward, A. (2000): Drinking water, ecology, and gastroenteritis in New Zealand. *IAHS-AISH Publication*, no. 260, pp. 41 - 47.

Weitzman, E. R.; Chen, Y. Y. & Subramanian, S. V. (2005): Youth smoking risk and community patterns of alcohol availability and control: a national multilevel study. *Journal of Epidemiology and Community Health 59*, no. 12, pp. 1065 - 1071.

Wheeler, J. G.; Sethi, D.; Cowden, J. M.; Wall, P. G.; Rodrigues, L. C.; Tompkins, D. S.; Hudson, M. J. & Roderick, P. J. (1999): Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. The Infectious Intestinal Disease Study Executive. *British Medical Journal 318*, no. 7190, pp. 1046 - 1050.

White, P. (2007): *correspondence via email, Christchurch, New Zealand. Date: 19/02/2007.*

WHO (1946): Constitution of the World Health Organization. Retrieved: 12/10/2006, from: http://w3.whosea.org/aboutsearo/pdf/const.pdf.

WHO (2000): Campylobacter. Retrieved: 08/08/2006, from: http://www.who.int/mediacentre/factsheets/fs255/en/.

WHO (2002): Foodborne diseases, emerging. Retrieved: 19/09/2006, from: http://www.who.int/mediacentre/factsheets/fs124/en/.

Wilkinson, M. & Flegg, K. M. (2001): Who are the GPs of New Zealand? Retrieved: 01/05/2007, from: http://www.rnzcgp.org.nz/news/nzfp/Aug2001/osp_wilkinson.htm.

Williams, C. J.; Willocks, L. J.; Lake, I. R. & Hunter, P. R. (2004): Geographic correlation between deprivation and risk of meningococcal disease: an ecological study. *BMC Public Health 4*, no. 30, pp. 1 - 11.

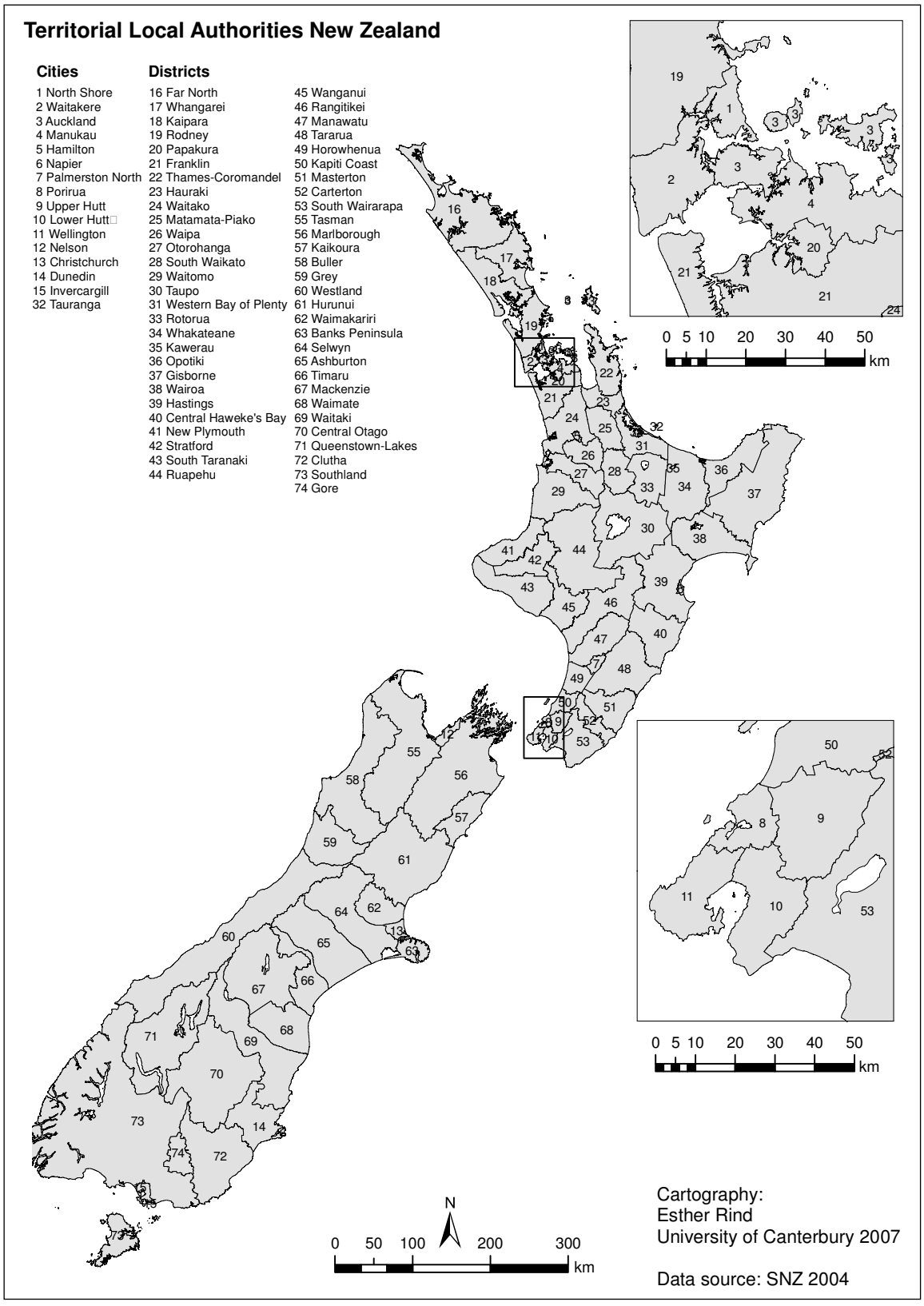
Wilson, N. (2005): A Systematic Review of the Aetiology of Human Campylobacteriosis in New Zealand. Retrieved: 08/08/2006, from: http://www.nzfsa.govt.nz/science/research-projects/campy-aetiol/campy-aetiol.pdf.

Winkler, E.; Turrell, G. & Patterson, C. (2006): Does living in a disadvantaged area mean fewer opportunities to purchase fresh fruit and vegetables in the area? Findings from the Brisbane food study. *Health and Place 12*, no. 3, pp. 306 - 319.

Wray, N. R.; Alexander, F. E.; Muirhead, C. R.; Pukkala, E.; Schmidtmann, I. & Stiller, C. (1999): A comparison of some simple methods to identify geographical areas with excess incidence of a rare disease such as childhood leukaemia. *Statistics in Medicine 18*, no. 12, pp. 1501 - 1516.

Appendix

A1: TLA-denotation map



A2: Description and sources of the original data

Variable	Description (units per TLA)	Type	Change of original data	Source
TAID	ID Territorial Local Authorities: 73 TLAs without Chatham Islands	spatial variables (shapefile)	no	GeoHealth Lab, Department of Geography, Canterbury NZ, original datasource: http://www.stats.govt.nz/default.htm
Name				
R9705	crude Campylobacter rate 1997 - 2005	Epidemiological data (dependent variable)	yes, aggregation	NZPHO 1997 - 2005, http://www.nzpho.org.nz/data_available.asp
Rain	monthly mean total rainfall (mm) - maximum	Climate: affects distribution of the organism	yes, aggregation	Climate Estimates from Landcare Research 1/5/2002, Ice-Server, Department of Geography, Canterbury NZ
Temp	monthly mean maximum daily temperatures (°C) - maximum	Climate: affects distribution of the organism	yes, aggregation	Climate Estimates from Landcare Research 1/5/2002, Ice-Server, Department of Geography, Canterbury NZ
WIS	weighted index of interseasonal change	Climate: affects distribution of the organism	yes - calculation	calculated from: Hearnden et al. 2003, The regionality of campylobacteriosis seasonality in New Zealand, Int. J. of Env. Health Research 13, 4, pp. 337 - 348
Pastoral_1	% of landuse primarily pastoral (high and low producing grassland)	landuse: affects distribution of the organism	yes - aggregation + conversion in %	New Zealand Landcover Database (LCDB2), Ministry of the Environment/ Terralink International/ ICE-server, Department of Geography, Canterbury NZ
Pastoral_2	% of landuse: scrub (e.g. fernland, gorse, broom, sub alpine shrubland); tussock grassland (e.g. low producing and depleted grassland)	landuse: affects distribution of the organism	yes - aggregation + conversion in %	New Zealand Landcover Database (LCDB2), Ministry of the Environment/ Terralink International/ ICE-server, Department of Geography, Canterbury NZ
Pastoral_3	total stock units per sqkm, including dairy, beef, sheep, deer, pigs, goats, horses, poultry	landuse: affects distribution of the organism	yes - SU-calculation	Statistics New Zealand/ MAF, table jointly compiled by StatsNZ and the Policy Information Group, Ministry of Agriculture and Forestry, http://www.maf.govt.nz/statistics/primary-industries/index.htm
ECEX	intensity of drinking water quality monitoring: % of E. coli exceedances in drinking water	exposure contaminated water	no	NZPHO 2002, http://www.nzpho.org.nz/data_available.asp
RW	% of population with rainwater supply	exposure contaminated water	yes - aggregation + conversion in %	WINZ 2006, http://www.drinkingwater.co.nz/supplies/supplies.asp and MoH publication: Register of Community Drinking-Water Supplies in NZ 2006
NAPWS	% of population not on registered water supply	exposure contaminated water	no	NZPHO 2002, http://www.nzpho.org.nz/data_available.asp
DWMM	number of drinking water measurements per capita	exposure contaminated water	no	NZPHO 2002, http://www.nzpho.org.nz/data_available.asp
Rurality_1	total stock units per average population 97_01, including dairy, beef, sheep, deer, pigs, goats, horses, poultry	exposure rurality, e.g. contact with carriers	yes - SU-calculation	Statistics New Zealand/ MAF, table jointly compiled by StatsNZ and the Policy Information Group, Ministry of Agriculture and Forestry, http://www.maf.govt.nz/statistics/primary-industries/index.htm
Rurality_2	% of population living in highly rural/ remote area	exposure rurality	yes - aggregation + conversion in %	Statistics NZ, 2001 Census of Population and Dwellings,http://www.stats.govt.nz/census/2001-census-data/default.htm
Occup	usual residence by occupation for the employed usually president population count, aged 15 years and over per 1,000 population working in dairy and meat related industries	exposure occupation	yes, aggregation	Statistics NZ, 2001 Census of Population and Dwellings, http://www.stats.govt.nz/census/2001-census-data/default.htm
Rest	Number of restaurants, bars, tavernen, clubs with alcohol licence per 100,000 population	exposure contaminated food	yes - aggregation	GeoHealth Lab, Department of Geography, Canterbury NZ
FaFo	number of fast food outlets per 100,000 population	exposure contaminated food	yes - aggregation	GeoHealth Lab, Department of Geography, Canterbury NZ
FreFo	number of fresh food outlets per 100,000 population	exposure contaminated food	yes - aggregation	GeoHealth Lab, Department of Geography, Canterbury NZ
SuMa	number of supermarkets per 100,000 population	exposure contaminated food	yes - aggregation	GeoHealth Lab, Department of Geography, Canterbury NZ
u15	% of people age < 15	demography, age	yes - conversion in %	Statistics NZ, 2001 Census of Population and Dwellings, http://www.stats.govt.nz/census/2001-census-data/default.htm
e25_44	% of people age 25 - 44	demography, age	yes - conversion in %	Statistics NZ, 2001 Census of Population and Dwellings, http://www.stats.govt.nz/census/2001-census-data/default.htm
male	% of males	demography, gender	yes - conversion in %	Statistics NZ, 2001 Census of Population and Dwellings, http://www.stats.govt.nz/census/2001-census-data/default.htm
Europ	% of Europeans	demography, ethnicity	yes - conversion in %	Statistics NZ, 2001 Census of Population and Dwellings, http://www.stats.govt.nz/census/2001-census-data/default.htm
NZDep	New Zealand Deprivation Score 2001	socioeconomic factors	yes - aggregation	GeoHealth Lab, Department of Geography, Canterbury NZ, calculation adapted from Salmond and Crampton 2001
GP	number of general practitioners per 100,000 population	notification artefact	yes - aggregation	GeoHealth Lab, Department of Geography, Canterbury NZ

A3: Data overview

TAID	Name	R9705	Rain	Temp	WIS	Pastoral_1	Pastoral_2	Pastoral_3	ECEX	RW	NAPWS	DWMM	Rurality_1	Rurality_2	Occup	Rest	FaFo	FreFo	SuMa	u15	e25_44	male	Europ	NZDep	GP
1	North Shore City	3397.68	154.00	24.10	3.54	18.77	7.17	1.59	0.00	0.00	0.00	0.39	0.00	0.00	1.65	123.50	90.68	97.91	12.24	20.66	30.83	48.25	74.82	938.96	42.72
2	Waitakere City	2512.12	239.00	24.10	8.25	16.89	6.58	106.06	0.10	0.23	3.90	0.71	0.24	0.00	2.88	47.68	70.91	82.52	9.78	24.74	32.64	48.81	61.22	1000.00	26.04
3	Auckland City	3029.79	265.00	24.40	9.73	17.66	43.49	23.64	0.20	0.83	8.50	0.49	0.04	0.00	1.66	255.89	114.01	135.75	8.36	19.70	34.44	48.40	57.75	996.41	51.99
4	Manukau City	2076.41	222.00	24.60	10.71	51.44	4.17	97.11	0.70	1.04	5.90	0.31	0.20	0.00	3.38	72.94	75.89	85.10	10.68	26.98	29.89	48.65	44.28	1044.30	31.08
5	Hamilton City	3404.18	147.00	24.70	20.30	35.52	0.56	79.74	0.00	0.00	4.00	0.28	0.07	0.00	8.34	129.18	89.98	96.22	14.25	22.60	29.58	47.93	69.82	1012.24	38.32
6	Napier City	2599.37	131.00	24.30	19.93	58.99	0.66	86.87	1.30	0.43	10.80	1.01	0.17	0.00	7.27	160.41	156.68	78.34	16.79	22.26	27.23	47.88	77.14	1018.20	48.41
7	Palmerston North City	1701.10	184.00	23.10	16.52	66.65	13.64	141.66	0.10	0.44	0.00	0.97	0.66	0.00	9.46	143.79	85.72	80.19	27.65	21.20	28.99	47.73	75.89	991.57	47.29
8	Porirua City	2632.03	199.00	21.70	23.42	51.94	24.36	43.49	0.00	0.11	6.60	1.62	0.16	0.01	7.27	65.90	70.15	63.78	31.89	28.06	30.51	49.07	53.77	1046.65	29.60
9	Upper Hutt City	3375.60	281.00	22.30	42.28	15.35	22.63	9.49	0.00	0.00	7.10	0.55	0.14	0.00	4.19	76.59	65.65	93.01	19.15	23.35	29.74	49.82	78.95	965.94	30.20
10	Lower Hutt City	3061.84	294.00	21.50	8.66	10.28	35.24	18.60	0.10	0.03	4.00	1.31	0.07	0.00	2.73	105.69	30.35	91.04	16.74	24.03	31.01	48.86	68.19	995.07	34.60
11	Wellington City	4735.11	172.00	21.10	12.03	37.33	38.77	27.87	0.00	0.04	5.30	3.68	0.05	0.00	1.90	238.60	149.36	113.41	17.35	18.53	36.17	48.38	73.42	952.75	30.53
12	Nelson City	1342.58	273.00	22.60	76.28	14.83	22.64	10.26	0.40	0.22	7.60	3.07	0.11	0.00	5.07	252.39	61.26	102.92	22.05	20.86	29.63	48.44	85.99	988.14	62.91
13	Christchurch City	3009.37	122.00	22.20	129.73	47.79	3.84	133.91	0.40	0.03	8.40	0.68	0.19	0.00	5.60	180.55	69.86	77.84	11.48	19.33	30.08	48.14	82.58	981.97	38.55
14	Dunedin City	2841.69	134.00	21.00	111.29	54.68	32.14	70.30	2.20	0.09	12.60	0.69	1.99	0.06	13.13	187.34	56.98	43.17	16.40	18.36	27.34	47.72	85.94	990.54	34.96
15	Invercargill City	2364.45	115.00	19.20	33.33	58.38	18.77	97.52	0.50	0.55	0.00	1.07	0.73	0.00	33.32	113.43	86.05	88.01	21.51	22.21	28.52	48.14	83.10	999.44	34.16
16	Far North District	969.42	297.00	25.00	17.78	42.16	17.70	44.06	3.50	12.54	38.90	1.27	5.48	3.70	41.49	248.77	70.55	155.95	31.56	26.33	25.93	49.53	51.96	1097.16	34.90
17	Whangarei District	2112.21	269.00	25.10	16.60	61.23	5.31	89.82	1.40	3.95	4.90	0.69	3.61	0.00	27.64	111.11	50.37	66.67	20.74	24.17	26.85	48.65	70.51	1028.13	39.65
18	Kaipara District	1512.15	306.00	24.80	34.81	70.80	5.91	84.50	2.10	9.55	45.10	1.60	15.13	31.41	111.53	195.67	80.57	143.87	23.02	24.88	25.75	50.10	73.29	1018.34	22.97
19	Rodney District	3077.33	207.00	24.10	88.12	63.01	8.61	147.26	2.30	5.83	52.80	1.08	4.84	0.00	21.51	173.04	66.45	148.12	18.00	22.71	28.06	48.95	83.75	956.78	28.88
20	Papakura District	2962.35	155.00	24.60	21.34	69.23	4.60	95.83	0.00	0.25	5.10	0.60	0.28	0.00	7.02	87.12	72.18	126.94	7.47	25.45	29.37	48.60	62.67	1031.28	52.67
21	Franklin District	3112.37	240.00	24.70	40.02	76.12	5.24	209.41	1.80	3.74	36.50	1.59	9.13	0.00	40.49	95.68	59.80	87.71	9.97	25.49	30.10	49.99	73.46	974.71	19.32
22	Thames-Coromandel District	2939.59	299.00	24.70	34.21	22.84	18.53	36.84	2.50	0.58	20.80	2.59	3.25	2.79	25.21	344.11	92.03	124.04	36.01	19.40	23.74	49.05	79.89	1017.72	43.66
23	Hauraki District	2853.62	328.00	25.10	38.63	65.32	4.32	140.98	2.50	0.39	9.80	1.85	9.85	14.46	87.36	100.01	70.59	141.18	29.41	25.39	25.91	49.62	77.74	1037.42	17.88
24	Waikato District	3272.96	247.00	24.90	56.54	76.97	3.07	153.07	1.30	1.91	44.90	1.84	12.14	0.00	75.43	70.83	40.47	73.35	27.82	26.89	28.56	49.98	68.38	1024.38	25.13
25	Matamata-Piako District	3859.35	287.00	25.20	67.12	84.00	0.52	340.05	1.40	0.51	41.40	1.92	20.21	0.00	128.68	84.70	60.99	125.36	27.10	24.94	28.05	49.61	81.85	980.93	20.38
26	Waipa District	3254.37	256.00	25.00	41.88	89.06	0.80	242.34	0.90	0.00	21.90	1.61	8.98	0.00	67.29	103.10	60.35	90.52	7.54	24.35	28.41	48.88	80.74	960.04	27.23
27	Otorohanga District	3025.51	258.00	24.90	34.52	66.40	5.62	130.40	1.10	0.14	35.50	2.78	27.60	0.00	156.01	137.69	42.36	116.50	10.59	25.89	29.80	53.70	68.37	938.30	21.47
28	South Waikato District	2201.20	231.00	25.00	92.10	36.30	1.34	119.62	0.60	0.00	4.40	2.18	9.04	0.00	54.04	83.21	99.85	128.97	29.12	28.77	28.30	50.23	58.88	1052.17	34.10
29	Waitomo District	2718.33	288.00	24.60	39.44	59.58	5.54	91.92	1.60	0.24	33.00	4.58	34.13	23.11	122.02	168.15	94.59	126.11	21.02	26.39	28.13	49.97	60.95	1033.05	31.89
30	Taupo District	2428.33	367.00	24.40	10.21	23.21	15.40	35.21	0.00	0.00	0.00	2.36	7.18	0.58	27.03	263.84	64.35	90.09	19.31	24.44	28.84	49.50	66.33	1023.46	25.45
31	Western Bay of Plenty District	2166.67	294.00	24.40	37.47	40.24	3.55	93.46	0.00	0.30	22.20	0.37	4.95	0.76	41.57	94.91	48.81	67.80	18.98	24.17	25.83	49.75	77.76	1008.06	20.93
32	Tauranga District	2187.41	178.00	23.90	9.81	37.94	4.98	54.94	0.20	0.00	6.90	0.54	0.09	0.00	5.68	170.69	71.32	104.05	16.37	21.67	27.28	47.64	78.31	1003.57	38.55
33	Rotorua District	2157.49	247.00	24.50	29.01	46.61	4.13	96.84	1.30	0.16	7.40	1.03	3.81	1.05	21.82	209.34	65.13	38.77	15.51	25.75	29.26	48.62	59.21	1029.07	34.09
34	Whakatane District	1500.55	250.00	25.10	21.07	18.85	2.14	36.15	2.00	0.15	4.30	13.49	4.88	0.00	37.67	100.33	51.68	72.97	24.32	27.22	27.50	49.13	56.92	1067.58	33.58
35	Kawerau District	1039.14	211.00	25.20	30.53	55.57	3.44	n/a	2.80	0.00	0.00	3.01	n/a	0.00	4.94	41.20	27.47	82.40	27.47	31.28	27.16	49.96	42.17	1135.87	43.10
36	Opotiki District	873.08	342.00	24.10	25.16	11.73	5.29	17.34	0.80	0.00	14.40	2.52	5.77	4.04	41.74	53.93	43.14	86.29	21.57	28.51	25.54	49.49	45.63	1146.96	54.29
37	Gisborne District	1782.65	450.00	24.30	20.52	48.82	16.15	65.96	4.80	5.55	24.40	0.79	12.33	6.26	27.52	129.75	85.01	136.47	6.71	27.48	27.35	48.92	53.24	1076.17	31.79
38	Wairoa District	1552.12	322.00	24.40	58.57	47.15	8.55	50.33	1.20	4.76	35.50	2.74	21.94	14.94	86.85	150.11	75.05	128.66	0.00	28.03	26.01	50.20	43.57	1091.48	67.23
39	Hastings District	3087.56	251.00	24.70	48.60	55.83	15.62	96.57	0.30	0.22	4.40	1.45	7.53	0.00	29.69	103.16	55.32	65.79	16.45	25.35	27.78	48.89	68.35	1027.59	29.64
40	Central Hawke's Bay District	2803.52	239.00	24.20	30.35	87.06	3.63	148.04	1.40	4.63	38.50	2.72	38.14	0.00	146.56	154.85	69.68	139.36	30.97	24.32	26.85	50.43	77.08	984.29	31.09
41	New Plymouth District	2530.29	541.00	23.20	20.68	52.70	5.65	290.18	0.60	0.75	20.20	1.02	9.52	0.00	35.60	132.36	80.31	62.46	23.80	22.81	27.43	48.40	81.82	1001.69	38.98
42	Stratford District	1995.88	560.00	24.50	40.37	43.86	8.35	77.28	0.00	2.93	31.70	1.11	18.33	1.92	126.60	153.45	76.73	153.45	10.96	24.48	27.42	49.46	87.25	987.29	33.90
43	South Taranaki District	2262.58	504.00	23.40	91.51	60.73	12.21	132.53	0.60	0.78	11.10	2.84	16.84	3.37	148.94	149.26	56.86	67.52	28.43	25.53	27.99	50.39	77.41	1004.14	50.94
44	Ruapehu District	1418.71	369.00	24.90	33.67	42.99	18.92	56.68	3.00	0.41	12.60	2.55	24.97	5.73	58.32	327.30	45.82	144.01	52.37	27.35	29.49	51.72	59.60	1057.30	55.82
45	Wanganui District	1793.60	228.00	24.60	23.10	50.96	19.56	94.59	3.40	1.10	27.90	0.78	5.11	0.00	28.51	95.71	102.55	107.10	9.12	23.32	25.84	48.05	74.20	1043.05	39.37
46	Rangitikei District	2224.98	271.00	23.70	5.07	61.77	23.00	107.49	1.60	2.38	29.90	1.99	30.89	0.20	81.59	179.61	51.32	89.80	38.49	25.25	27.52	49.88	73.83	1008.88	26.48
47	Manawatu District	1374.50	203.00	23.40	9.34	82.02	10.00	152.08	1.70	5.08	33.20	1.03	14.40	0.00	72.72	115.44	68.54	64.93	10.82	24.53	28.53	49.62	83.11	975.78	36.42
48	Tararua District	1337.56	333.00	22.90	10.37	82.85	11.96	137.32	3.30	5.38	27.10	2.03	32.63	3.38	137.81	179.82	43.59	59.94	38.14	25.61	27.33	50.31	79.38	995.12	33.48
49	Horowhenua District	1466.45	331.00	22.90	6.37	61.65	9.83	161.90	0.90	1.27	12.40	1.13	5.76	0.00	32.79	93.58	93.58	96.92	20.05	22.71	24.88	48.48	74.84	1044.78	33.58
50	Kapiti Coast District	2922.48	286.00	22.00	36.84	28.72	11.44	29.19	0.00	0.00	8.70	2.43	0.52	0.02	7.41	146.74	83.15	66.03	29.35	20.96	23.94	46.90	82.67	968.64	37.66
51	Masterton District	1101.86	344.00	23.40	5.77	70.57	11.82	112.35	3.90	0.57	7.90	1.11	11.39	0.00	32.02	101.43	70.56	105.84	17.64	23.39	25.60	48.47	79.56	1004.78	39.71
52	Carterton District	2014.45	337.00	23.40	25.57	60.86	15.18	88.77	0.00	0.00	36.40	0.98	15.40	0.00	73.64	161.67	88.19	146.98	14.70	22.76	26.65	49.12	85.66	975.92	14.58
53	South Wairarapa District	4061.32	272.00	23.40	3.63	51.62	23.81	73.81	2.40	0.61	24.60	2.37	20.08	0.00	74.00	409.22	68.20	318.28	56.84	22.41	25.98	49.35	82.31	980.53	34.32
55	Tasman District	1515.27	475.00	23.70	20.72	15.30	15.57	18.24	1.10	0.39	42.10	2.11	4.40	2.84	37.09	219.81	92.42	134.88	27.48	22.97	28.68	49.98	87.98	968.02	33.91
56	Marlborough District	2037.54	420.00	23.70	170.80	20.57	40.44	22.83	5.30	0.12	16.90	1.80	6.12	0.02	36.13	250.61	127.86	191.79	25.57	20.96	26.85	49.56	85.84	967.78	32.89
57	Kaikoura District	2364.67	267.00	22.50	159.58	14.53	62.98	29.45	2.50	0.00	0.00	4.54	17.48	0.00	66.13	1044.20	174.03	232.05	58.01	20.00	28.10	49.61	80.94	980.17	28.89
58	Buller District	2295.08	680.00	23.00	145.48	6.18	16.84	8.05	1.20	2.67	18.90	2.52	6.44	9.95	46.78	452.66	80.47	171.01	50.30	22.49	27.48	50.73	87.04	1043.04	41.38
59	Grey District	2006.31	569.00	22.40	20.07	11.99	19.11	12.42	0.60	3.82	12.50	4.06	3.32	2.58	28.99	304.35	98.91	182.61	15.22	23.18	29.19	50.02	88.78	1003.71	39.02
60	Westland District	2256.30	1096.00	21.10	41.38	5.56	25.43	5.56	6.50	1.75	18.40	0.38	8.39	30.05	66.99	735.87	76.12	101.50	50.75	21.33	29.81	50.42	83.32	991.39	51.70
61	Hurunui District	3944.77	478.00	23.70	64.98	33.12	41.08	52.62	8.50	0.26	0.00	6.78	47.20	59.37	137.34	424.65	41.43	227.86	72.50	23.30	27.21	50.47	90.62	961.07	50.67
62	Waimakariri District	3054.85	176.00	22.70	118.11	53.86	20.15	177.32	0.40	0.00	27.60	1.36	11.24	0.00	41.82	91.40	45.70	79.98	14.28	23.61	29.52	49.34	89.93	946.90	38.05
63	Banks Peninsula District	3712.36	236.00	22.40	64.22	75.94	19.40	78.62	1.80	0.00	8.00	8.07	10.03	21.34	39.86	0.00	52.11	169.35	13.03	19.34	30.87	49.85	87.39	941.88	38.42
64	Selwyn District	4160.19	841.00	22.90	105.82	36.04	27.31	87.53	1.90	0.00	39.10	2.39	21.33	0.00	78.81	0.00	37.96	113.89	26.57	23.60	30.99	51.50	89.05	913.57	14.64
65	Ashburton District	3591.32	900.00	22.90	183.28	50.47	26.58	92.16	1.70	0.30	7.30	2.23	22.56	0.00	118.91	0.00	67.25	90.99	27.69	20.92	27.38	50.12	92.75	943.58	31.49
66	Timaru District	3891.82	608.00	22.10	160.10	56.57	25.88	101.08	2.40	0.00	0.00	2.02	6.55	0.00	53.53	0.00	73.46	104.26	21.33	20.96	26.29	48.42	90.52	975.60	50.21
67	Mackenzie District	3415.51	808.00	23.10	47.85	13.76	54.96	22.89	12.80	0.00	0.00	7.49	43.35	72.76	116.91	0.00	103.23	154.85	77.42	20.65	29.88	51.49	87.18	953.45	53.25
68	Waimate District	2673.18	142.00	23.50	183.60	49.12	45.44	67.14	13.30	0.00	13.10	4.52	32.98	0.00	151.38	0.00	54.85	82.27	13.71	21.76	24.84	50.82	93.05	980.86	42.30
69	Waitaki District	2231.34	573.00	23.50	144.20	42.85	43.48	45.68	9.80	0.00	0.60	3.99	15.82	6.01	88.91	0.00	72.64	150.11	48.42	19.94	24.46	49.00	91.35	974.48	44.90
70	Central Otago District	2196.61	182.00	24.40	76.32	40.82	54.77	40.81	1.90	0.00	16.10	3.13	27.74	24.06	58.77	0.00	81.90	184.29	75.08	19.84	24.85	50.28	90.85	955.25	41.49
71	Queenstown-Lakes District	4125.55	672.00	23.90	98.72	7.45	57.47	12.02	0.00	0.10	7.80	3.29	6.58	2.34	15.67	0.00	132.18	195.12	44.06	18.12	36.73	49.97	85.14	931.46	41.09
72	Clutha District	3774.76	218.00	21.60	61.12	66.53	12.46	106.38	2.70	3.02	8.40	5.59	38.67	37.59	162.32	0.00	5.71	39.99	57.13	23.18	28.69	51.64	89.00	955.83	52.44
73	Southland District	2914.24	819.00	22.00	285.37	26.77	25.79	48.66	5.70	2.63	52.60	3.38	49.12	26.64	188.19	0.00	54.43	142.87	30.62	23.28	29.28	52.23	88.44	941.81	24.44
74	Gore District	3229.07	141.00	21.10	144.13	96.64	4.22	210.97	0.00	0.00	14.00	1.71	20.72	0.00	130.67	0.00	86.33	102.02	31.39	22.62	28.16	49.67	88.06	965.07	32.08

A4: Calculation of the weighted index of interseasonal change

TAID	Name	SON	SP_SU	DJF	SU_A	MAM	A_W	JJA	W_SP	C_agg	group	WIS
1	North Shore City	318	7.23	341	-17.30	282	-4.61	269	18.22	3.54	1	3.54
2	Waitakere City	248	1.61	252	-26.19	186	1.61	189	31.22	8.25	1	8.25
3	Auckland City	286	12.94	323	-18.27	264	-9.47	239	19.67	4.87	2	9.73
4	Manukau City	204	18.63	242	-29.75	170	-6.47	159	28.30	10.71	1	10.71
5	Hamilton City	337	2.08	344	-31.98	234	17.95	276	22.10	10.15	2	20.30
6	Napier City	219	18.26	259	-23.94	197	-14.72	168	30.36	9.96	2	19.93
7	Palmerston North City	174	12.64	196	-21.94	153	-24.18	116	50.00	16.52	1	16.52
8	Porirua City	308	-2.60	300	-30.00	210	6.19	223	38.12	11.71	2	23.42
9	Upper Hutt City	276	44.20	398	-26.38	293	-16.38	245	12.65	14.09	3	42.28
10	Lower Hutt City	255	21.18	309	-19.74	248	-1.61	244	4.51	4.33	2	8.66
11	Wellington City	439	11.62	490	-19.39	395	-4.05	379	15.83	4.01	3	12.03
12	Nelson City	68	148.53	169	-41.42	99	-32.32	67	1.49	76.28	1	76.28
13	Christchurch City	234	83.76	430	-29.53	303	-40.26	181	29.28	43.24	3	129.73
14	Dunedin City	305	30.16	397	-19.40	320	-44.06	179	70.39	37.10	3	111.29
15	Invercargill City	146	39.04	203	-11.33	180	-25.56	134	8.96	11.11	3	33.33
16	Far North District	87	4.60	91	-29.67	64	-12.50	56	55.36	17.78	1	17.78
17	Whangarei District	219	-10.96	195	-18.46	159	-13.84	137	59.85	16.60	1	16.60
18	Kaipara District	246	-1.63	242	-47.93	126	15.87	146	68.49	34.81	1	34.81
19	Rodney District	235	33.62	314	-54.14	144	-2.08	141	66.67	44.06	2	88.12
20	Papakura District	309	15.86	358	-36.87	226	-10.62	202	52.97	21.34	1	21.34
21	Franklin District	321	-4.36	307	-37.46	192	10.42	212	51.42	20.01	2	40.02
22	Thames-Coromandel District	317	-7.26	294	-33.33	196	8.16	212	49.53	17.10	2	34.21
23	Hauraki District	381	-46.19	205	-13.17	178	20.79	215	77.21	38.63	1	38.63
24	Waikato District	319	-28.53	228	-42.98	130	15.38	150	112.67	56.54	1	56.54
25	Matamata-Piako District	634	-39.59	383	-44.91	211	46.45	309	105.18	67.12	1	67.12
26	Waipa District	329	-33.43	219	-38.81	134	46.27	196	67.86	41.88	1	41.88
27	Otorohanga District	419	-32.70	282	-34.40	185	41.08	261	60.54	34.52	1	34.52
28	South Waikato District	322	-40.99	190	-41.05	112	8.04	121	166.12	92.10	1	92.10
29	Waitomo District	304	25.33	381	-41.47	223	1.35	226	34.51	19.72	2	39.44
30	Taupo District	184	-32.61	124	12.90	140	6.43	149	23.49	10.21	1	10.21
31	Western Bay of Plenty District	201	-9.95	181	-42.54	104	3.85	108	86.11	37.47	1	37.47
32	Tauranga District	158	25.95	199	-28.14	143	-7.69	132	19.70	9.81	1	9.81
33	Rotorua District	153	-18.30	125	-19.20	101	-17.82	83	84.34	29.01	1	29.01
34	Whakatane District	190	-11.58	168	-38.10	104	27.88	133	42.86	21.07	1	21.07
35	Kawerau District	197	-48.22	102	50.00	153	0.00	153	28.76	30.53	1	30.53
36	Opotiki District	271	-34.69	177	-23.16	136	31.62	179	51.40	25.16	1	25.16
37	Gisborne District	161	1.24	163	-40.49	97	13.40	110	46.36	20.52	1	20.52
38	Wairoa District	162	62.35	263	-10.27	236	-42.80	135	20.00	29.28	2	58.57
39	Hastings District	254	27.17	323	-39.94	194	49.48	290	-12.41	24.30	2	48.60
40	Central Hawke's Bay District	178	35.96	242	-43.80	136	52.21	207	-14.01	30.35	1	30.35
41	New Plymouth District	203	-7.88	187	-39.57	113	30.97	148	37.16	20.68	1	20.68
42	Stratford District	371	-29.11	263	-40.30	157	73.89	273	35.90	40.37	1	40.37
43	South Taranaki District	278	-41.37	163	-25.15	122	-17.21	101	175.25	91.51	1	91.51
44	Ruapehu District	158	56.33	247	2.02	252	-23.81	192	-17.71	16.84	2	33.67
45	Wanganui District	147	-2.72	143	-42.66	82	43.90	118	24.58	23.10	1	23.10
46	Rangitikei District	196	-22.45	152	19.74	182	8.79	198	-1.01	5.07	1	5.07
47	Manawatu District	131	16.03	152	-28.95	108	-3.70	104	25.96	9.34	1	9.34
48	Tararua District	162	-11.73	143	-22.38	111	3.60	115	40.87	10.37	1	10.37
49	Horowhenua District	146	11.64	163	-19.02	132	-12.12	116	25.86	6.37	1	6.37
50	Kapiti Coast District	241	47.30	355	-30.99	245	20.41	295	-18.31	18.42	2	36.84
51	Masterton District	95	14.74	109	-17.43	90	-13.33	78	21.79	5.77	1	5.77
52	Carterton District	359	-28.13	258	-15.89	217	-5.53	205	75.12	25.57	1	25.57
53	South Wairarapa District	345	4.06	359	-13.93	309	11.97	346	-0.29	1.82	2	3.63
55	Tasman District	122	44.26	176	-37.50	110	-11.82	97	25.77	20.72	1	20.72
56	Marlborough District	142	124.65	319	-52.98	150	-37.33	94	51.06	85.40	2	170.80
57	Kaikoura District	739	-27.60	535	0.93	540	-36.85	341	116.72	53.19	3	159.58
58	Buller District	312	-29.81	219	4.11	228	-43.42	129	141.86	72.74	2	145.48
59	Grey District	249	-4.02	239	-38.08	148	12.16	166	50.00	20.07	1	20.07
60	Westland District	322	3.73	334	-17.66	275	-28.00	198	62.63	20.69	2	41.38
61	Hurunui District	517	-5.03	491	-31.36	337	-7.12	313	65.18	21.66	3	64.98
62	Waimakariri District	254	48.43	377	-28.12	271	-41.70	158	60.76	39.37	3	118.11
63	Banks Peninsula District	290	48.28	430	-38.84	263	-8.37	241	20.33	21.41	3	64.22
64	Selwyn District	347	34.29	466	-31.55	319	-35.11	207	67.63	35.27	3	105.82
65	Ashburton District	316	30.70	413	-26.63	303	-49.50	153	106.54	61.09	3	183.28
66	Timaru District	266	83.46	488	-36.07	312	-42.63	179	48.60	53.37	3	160.10
67	Mackenzie District	575	-4.17	551	26.13	695	-36.69	440	30.68	15.95	3	47.85
68	Waimate District	365	52.33	556	-35.43	359	-46.80	191	91.10	61.20	3	183.60
69	Waitaki District	246	-7.72	227	27.75	290	-60.00	116	112.07	72.10	2	144.20
70	Central Otago District	162	78.40	289	-48.10	150	2.67	154	5.19	38.16	2	76.32
71	Queenstown-Lakes District	373	38.87	518	-41.31	304	-22.04	237	57.38	32.91	3	98.72
72	Clutha District	203	40.39	285	-17.89	234	-42.31	135	50.37	30.56	2	61.12
73	Southland District	301	1.66	306	-27.78	221	-58.37	92	227.17	142.69	2	285.37
74	Gore District	202	68.32	340	-27.94	245	-46.53	131	54.20	48.04	3	144.13

TAID: TLA identification number
SON: aggregated spring rate 1993 - 2000
SP_SU: seasonal change from spring to summer
DJF: aggregated spring summer rate 1993 - 2000
SU_A: seasonal change from summer to autumn
MAM: aggregated autumn rate 1993 - 2000
A_W: seasonal change from autumn to winter
JJA: aggregated winter rate 1993 - 2000
W_SP: seasonal change from winter to spring
C_agg: aggregated interannual change
group: classes seasonal grouping
WIS: weighted index of interseasonal change

Original data and grouping: Hearnden et al. 2003

A5: Correlation matrix of the independent variables

	Rain	Temp	WIS	Pastoral_1	Pastoral_2	Pastoral_3	ECEX	RW	NAPWS	DWMM	Rurality_1	Rurality_2	Occup	Rest	FaFo	FreFo	SuMa	u15	e25_44	male	Europ	NZDep	GP
Rain	1.000																						
Temp	-0.160	1.000																					
WIS	.359(*)	-.288(**)	1.000																				
Pastoral_1	-.392(*)	0.126	-0.090	1.000																			
Pastoral_2	.329(*)	-.357(C11*)	.390(*)	-.461(*)	1.000																		
Pastoral_3	-0.230	0.173	-0.027	.771(*)	-.499(*)	1.000																	
ECEX	.360(*)	-0.021	.355(*)	-0.097	.477(*)	-0.172	1.000																
RW	0.037	0.181	-0.137	0.224	-0.186	0.052	0.071	1.000															
NAPWS	0.116	.231(**)	0.082	.310(*)	-.240(**)	.316(*)	-0.010	.542(*)	1.000														
DWMM	0.141	-0.015	0.175	-0.161	.275(**)	-0.218	.385(*)	-0.124	-0.153	1.000													
Rurality_1	.305(*)	0.026	.337(*)	0.230	.238(**)	0.127	.546(*)	0.132	.330(*)	.420(*)	1.000												
Rurality_2	.379(*)	-0.085	0.108	-0.139	.312(*)	-0.207	.571(*)	0.109	-0.013	.460(*)	.600(*)	1.000											
Occup	.323(*)	0.005	.402(*)	.343(*)	0.056	.265(**)	.432(*)	0.195	.416(*)	.327(*)	.886(*)	.445(*)	1.000										
Rest	0.134	-0.102	-0.030	-.348(*)	.246(**)	-.284(**)	0.040	0.088	-0.033	-0.001	-0.052	0.058	-0.068	1.000									
FaFo	0.009	-0.064	0.033	-.275(**)	.331(*)	-.232(**)	-0.053	-0.092	-0.159	-0.086	-.276(**)	-0.088	-.281(**)	.372(*)	1.000								
FreFo	.246(**)	0.097	0.170	-.248(**)	.459(*)	-.287(**)	.243(**)	0.086	0.132	.238(**)	.271(**)	.284(**)	0.166	.410(*)	.360(*)	1.000							
SuMa	.361(*)	-0.149	0.203	-0.223	.525(*)	-0.221	.447(*)	-0.059	-0.137	.401(*)	.516(*)	.596(*)	.346(*)	.307(*)	0.020	.450(*)	1.000						
u15	-0.139	.480(*)	-.309(*)	.250(**)	-.563(*)	.240(**)	-0.125	.265(**)	0.230	-0.024	0.123	-0.089	0.134	-0.171	-.439(*)	-0.230	-0.203	1.000					
e25_44	0.058	-0.200	-0.073	-.242(**)	0.201	-0.115	-.277(**)	-0.191	-0.163	-0.001	-0.185	-0.027	-.249(**)	-0.015	0.179	-0.026	-0.136	-0.170	1.000				
male	.417(*)	0.099	.316(*)	-0.006	0.169	-0.034	.358(*)	0.153	.373(*)	.365(*)	.726(*)	.440(*)	.714(*)	0.001	-.322(*)	.253(**)	.396(*)	.264(**)	0.067	1.000			
Europ	.316(*)	-.458(C24*)	.477(*)	0.027	.376(*)	0.086	.259(**)	-0.167	0.077	0.103	.273(**)	0.194	.326(*)	0.037	0.078	0.219	.325(*)	-.712(*)	-0.094	0.102	1.000		
NZDep	-0.208	.394(*)	-.360(*)	-0.033	-.389(*)	-0.094	-0.097	.238(**)	-0.089	-0.068	-.261(**)	-0.147	-.260(**)	0.058	-0.097	-0.171	-0.170	.678(*)	-.331(*)	-0.169	-.791(*)	1.000	
GP	0.071	-0.102	-0.008	-.268(**)	0.193	-.344(*)	0.214	-0.077	-.406(*)	0.159	-0.008	.275(**)	-0.111	0.099	0.039	0.017	0.206	-0.140	-0.034	-0.097	-0.076	0.215	1.000

Pearson correlation coefficient: * correlation is significant at the 0.01 level; ** correlation is significant at the 0.05 level; grey background indicates multicollinearity.